DOCUMENT RESUME

ED 278 841 CE 046 389

AUTHOR	Simpson, Carl
TITLE	JTPA Evaluation at the State and Local Level. Volume IV: A Guide for Gross Impact Evaluations.
INSTITUTION	Washington State Dept. of Employment Security, Olympia.
SPONS AGENCY	International Business Machines Corp., Armonk, N.Y.; National Commission for Employment Policy (DOL), Washington, D.C.
PUB DATE	Mar 86
NOTE	243p.; A product of the JTPA Evaluation Design Project. For related evaluation materials, see CE 046 385-393.
PUB TYPE	Guides - Non-Classroom Use (055)
EDRS PRICE	MF01/PC10 Plus Postage.
DESCRIPTORS	Data Analysis; Data Collection; Educational Legislation; *Employment Programs; Evaluation Criteria; *Evaluation Methods; Federal Legislation; Followup Studies; Literature Reviews; Local Issues; Measurement Techniques; Outcomes of Education; *Program Evaluation; *Research Design; Research Methodology; *Statewide Planning; Surveys
IDENTIFIERS	*Impact Studies; *Job Training Partnership Act 1982; Service Delivery Areas

ABSTRACT

This guide is intended to assist states and service delivery areas (SDAs) in addressing the new oversight responsibilities and opportunities stipulated by the Job Training Partnership Act (JTPA) with respect to planning and conducting a gross impact evaluation. The following topics are covered in the individual chapters: introduction and overview (reasons for studying gross impact, a quick map to gross impact analysis, and issues underlying gross impact analysis); research design (designs for participant and employer follow-ups and summaries of recommended employer and participant designs); measurement (issues underlying interpretations of JTPA intervention, participant measures, and employer measures); and analysis procedures (analyzing various types of measures, estimating change, analyzing a large number of potential influences on outcomes, and combining data from different sources). The section of research design for a participant follow-up contains detailed guidelines pertaining to demarcating alternative levels of investment in the analysis of participant outcomes, identifying the population to be analyzed, designing and selecting the sample, deciding what follow-up period to use, and using different data collection methods. Appendixes include a review of selected literature, background to gross impact analysis methods, information on selected implementation topics, illustrative applications, and selected measurement instruments. (MN)

****
*   Reproductions supplied by EDRS are the best that can be made   *
*   from the original document.   *
****

ED278841

JTPA
EVALUATION
DESIGN PROJECT

JTPA Evaluation at the State and Local Level

Volume IV: A Guide for Gross Impact Evaluations

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it
☐ Minor changes have been made to improve reproduction quality

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

D Feldman

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CE046389

BEST COPY AVAILABLE

By Carl Simpson

March 1986

BEST COPY AVAILABLE

Special appreciation is expressed to the National Commission for Employment Policy, for serving as the project's national sponsor and contributing substantial staff consultation to the project as it developed.

Project Development and Coordination:

Washington State Employment Security
Isiah Turner, *Commissioner*

Project Funding:

National Commission for Employment Policy
IBM Corporation
Washington State Employment Security

CONTEXT OF THIS VOLUME

This is one in a series of volumes produced by the JTPA EVALUATION DESIGN PROJECT.

PURPOSE AND PHILOSOPHY

The purpose of this project has been to develop a set of evaluation tools that are useful to states and local service delivery areas (SDAs) in judging the way their JTPA programs are being managed and the impact they are having. The intention has been to base these analytic and managerial tools on sound program concepts and research methods, and to design them such that the information obtained is of practical and direct use in improving JTPA policies and programs at the state and local level. This kind of information is also expected to make a unique contribution to national training policy and Federal oversight of JTPA.

It is hoped that these volumes will stimulate and support state and local evaluation efforts in JTPA, and promote more consistency than in previous programs with respect to the issues studied and the methods used to investigate them. An important goal is to encourage the generation of complementary information on program implementation and impact that is comparable across states and SDAs. Comprehensive, comparable information is essential to the development of a valid and reliable knowledge base for resolving problems and improving programs. It is also required for adjusting national training strategies to changing needs and priorities at the state and local level.

PRODUCTS

Consistent with this purpose and philosophy, the project has produced a set of materials to assist states and SDAs in evaluating their programs. These are to be useful in planning, designing and implementing evaluation activities. As an integrated collection, each set is developed to support comprehensive evaluations over the JTPA planning cycle.

The careful tailoring of these materials to state and local users is appropriate. JTPA represents a new employment and training policy shaped not only by the experience of managers and the perspectives of employers, but by scientific assessments of previous approaches for addressing unemployment, poverty and other barriers to economic security. In this context, the value of JTPA programs is also expected to be judged. In fact, the Act's assessment requirements are more explicit and sophisticated than those of any employment and training legislation to date. It clearly distinguishes between *monitoring* activities, whose purpose is to determine compliance (such as with performance standards) and *evaluation* activities, whose purpose is to determine how a program is being managed and implemented, and the kinds of effects it is having on recipients and relevant others. Equally significant, new constitutencies are expected to make these more rigorous assessments. States and SDAs now have this important responsibility. It is the first time in the history of employment and training programs that the Federal government's evaluation role has been significantly reduced.

This change affords states and local areas opportunities to influence public policy. It also requires them to assume new oversight responsibilities. Program evaluation is expected to become an integral part of the management of organizations administering, planning and delivering public training services. This is as it should be. The more information available at these levels, where changes in organizations can most readily be made, the more effective the management of JTPA programs. This project was undertaken in that context.

The evaluation tools produced by the project have been developed with a sensitivity to the differing needs, interests and resources of state and local users. They have been packaged into a single comprehensive and integrated set of volumes called *JTPA Evaluation at the State and Local Level*. The set contains planning and evaluation *guides* and *issue papers*. The following volumes are available in the set:

Volume	Author
I: Overview	Project Team
II: A General Planning Guide	Deborah Feldman
III: A Guide for Process Evaluations	David Grembowski
III Supplement: Some Process Issues at the State Level	David Grembowski
IV: A Guide for Gross Impact Evaluations	Carl Simpson
V: A Guide for Net Impact Evaluations	Terry Johnson
VI: An Implementation Manual for Net Impact Evaluations	Terry Johnson
VII: Issues Related to Net Impact Evaluations	
A. Issues in Evaluating Costs and Benefits	Ernst Stromsdorfer
B. The Debate Over Experimental vs. Quasi-Experimental Approaches	Ann Blalock
VIII: MIS Issues in Evaluating JTPA	David Grembowski

NOTE: Although each of the discrete products listed above is the responsibility of a single author, each seeks to incorporate the results of professional peer review, the many excellent recommendations of the advisory group, and the ideas and suggestions of the numerous practitioners interviewed in the process of developing these materials.

To further qualify these volumes, Volume III is accompanied by a supplement for state users. This is consistent with the significant differences between states and SDAs in the kinds of process issues that are most essential to study. The volume on net impact evaluations is sufficiently technical, because of the statistical methods involved, that a practical manual has been written to accompany it. This guide and manual tend to be more appropriate for states, since relatively large sample sizes are required for analysis. However, they are equally useful to larger SDAs and consortia of smaller SDAs which may want to jointly study the net impact of their programs. Regional evaluations, for example, can be very productive in providing management information relevant to regional labor markets. Although there is a separate issue paper on evaluating costs and benefits, this issue is also covered in the gross impact and net impact guides. In this respect, the user benefits from three related but different approaches to this important element of program evaluations. Also, the user should be aware that the Appendix of Volume II includes *A Report on a National/State Survey of Local JTPA Constituencies*. This survey was carried out by Bonnie Snedeker, with the assistance of Brian O'Sullivan, to provide additional input from practitioners to the development of the planning and process evaluation guides.

In conclusion, several expectations have directed the development of these volumes:

THE GUIDES

The General Planning Guide

This guide is to assist users in *planning, funding and developing an organizational capacity* to carry out process, gross outcome, and net impact evaluations and to *utilize their results*. Separate state and local versions are available.

The Evaluation Guides

These volumes are to have the following characteristics:

☐ The guides are to *complement* one another.

- They are to provide information on program management and other characteristics of program implementation, which can:
  - —Describe the way in which administrative, managerial and service delivery policies and practices operate to affect outcomes, as a set of interventions separate from the program's services.
  - —Pinpoint the source, nature and extent of errors and biases for which adjustments must be made in gross and net impact evaluations.
  - —Help explain the results of gross and net impact evaluations.
- They are to provide information on aggregate gross outcomes, and outcomes differentiated by type of service and type of recipient, which can:
  - —Describe relationships between certain implementation modes and service strategies, and a broad array of client and employer outcomes.
  - —Help explain the results of net impact evaluations.
  - —Suggest the more important outcomes that should be studied in net impact evaluations.
  - —Help sort out those aspects of implementation that may be most critical to study in process evaluations.
- They are to provide information on net impact (the program's return on investment), which can:
  - —Closely estimate the effect of the program's services on clients.
  - —Suggest which services and client groups are most important to study in broader but less rigorous gross impact studies.
  - —Help identify the decision points in program implementation (particularly service delivery) which may be most important to study in process evaluations.

☐ The guides are to enable the user to carry out *comprehensive assessments* of JTPA programs.

- They are to allow the user to acquire several different perspectives on the same program within a particular time period: on program implementation, on outcomes for clients and employers and on net impact.
- They are to permit the user to interrelate these different kinds of information to gain a wider understanding of what is happening in a program and why.

☐ The guides are to describe approaches and methodologies as consistently as possible, to achieve *comparability*.

- They are to define variables and relationships as similarly as possible.
- They are to define research designs, and methods of data collection and analysis using as similar concepts as possible.

☐ The guides are to *draw from past research* on employment and training programs, as well as seek *new* approaches and methods of specific value in evaluating JTPA at the state and local level.

- They are to replicate, to the extent possible and feasible, the issues and measures reflected in Federal monitoring and evaluation decisions.
- They are to make selective use of the results of relevant CETA studies, national studies of JTPA, and issue papers on JTPA evaluation by national public interest organizations in the employment and training area.
- They are to rely on the professional literature in applied social research.

THE ISSUE PAPERS

Volume VII contains two issue papers which serve as companion pieces to the preceding volumes on net impact evaluation. The first paper on cost-benefit issues is designed to help users identify, measure and analyze relationships between monetary and nonmonetary costs and benefits in determining the program's return on investment. The second paper examines the pros and cons of different research strategies associated with the net impact approach. The final volume on MIS issues is to assist users in better understanding how JTPA and other employment and training *management information systems* can efficiently support the evaluation of program implementation and impact.

THE SET OF VOLUMES

The set is *integrated*, but affords *flexible use*. The user can utilize the entire set for comprehensive evaluations over a two-year planning cycle or longer planning period, or the user can apply the information in each volume independently, based on the most pressing evaluation priorities and timeframes and given the extent of resources, during a particular fiscal year or biennium.

It should be understood that although evaluation products have been developed for JTPA, their basic principles and methods can be applied more broadly by states and local areas to evaluate other employment and training programs and other social programs.

GENERAL ACKNOWLEDGMENTS

The JTPA EVALUATION DESIGN PROJECT was developed and carried out based on the partnership philosophy that underlies the JTPA legislation. Several partnerships should be recognized for their substantial contributions to the products previewed here: the project development and coordination partnership; the public-private funding partnership; the interdisciplinary design partnership; and the advisory partnership.

The Development and Coordination Partnership: Washington Employment Security Department

Project Coordinators: Gary Bodeutsch, Ann Blalock

Assistant Coordinators: Deborah Feldman, Chris Webster

Interdivisional Consultants:

Training Division: Martin McCallum, Brian O'Neall, Ross Wiggins

Research and Information Systems: Jeff Jaksich, Mike Gioimo, Irv Lefberg

The Public/Private Funding Partnership

Funders:

IBM Corporation: Corporate Support Program, Armonk, NY
National Commission for Employment Policy, Washington, D.C.
Washington State Employment Security Department, Olympia, WA

Contributors:

Safeco Insurance Company of America, Seattle, WA
Seattle/King County Private Industry Council, Seattle, WA
SPSS, INC., Chicago, IL

The Interdisciplinary Design Partnership

Consultant:	*Area of Expertise:*	*Affiliation:*
Burt Barnow	Labor Economics	ICF, INC., Washington, D.C.
David Grembowski	Planning	University of Washington, Seattle, WA
Terry Johnson	Labor Economics	Battelle Memorial Institute, Seattle, WA
Brian O'Sullivan	Planning	Seattle/King County Private Industry Council, Seattle, WA
Carl Simpson	Sociology	Western Washington University, Bellingham, WA
Bonnie and David Snedeker	Management	Snedeker Scientific, Inc., Seattle, WA
Ernst Stromsdorfer	Labor Economics	Washington State University, Pullman, WA

The Advisory Partnership

The National Advisory Committee:

State of California: Tim Taormina
San Francisco SDA: Greg Marutani

State of Kansas: Richard Hernandez
Pittsburg SDA: John Gobetz

State of Massachusetts: Kay Stratton
New Bedford SDA: Paul Vigeant

State of New York: Robert Bernstein
Dunkirk PIC: Bruce Ritenberg III

State of Texas: Rick Mackay
Fort Worth PIC: Judge Michael Moncrief

State of Washington: Ernest La Palm
Seattle/King County PIC: Al Starr

Additional Advisors to the Project:

National Governors Association: Kay Albright, Washington, D.C.

HRDI: AFL-CIO: Candace Brown, Helena, MT

U.S. Department of Labor: Hugh Davies, Washington, D.C.

National Association of Counties: Kathy DuChene, Washington D.C.

State of Ohio Impact Assessment Unit: Gail James, Columbus, OH

Berkeley Planning Associates: Deborah Kogan, Berkeley, CA

Abt Associates: Jeff Zornitsky, Cambridge, MA

Special appreciation is expressed to the Commissioner and Deputy Commissioner of the Employment Security Department, Isiah Turner and Ernest La Palm, for supporting the need for evaluation at the state and local level and intitiating this project to address that need. They provided the strong encouragement and the resources necessary for designing the project, seeking a funding base for carrying it out and coordinating its development.

Other individuals made unique contributions to the development of the project: John Wallace and Ann Donohue of the National Commission for Employment Policy; Kay Albright of the National Governors Association; Dan Kelly of IBM's Corporate Support Program and Jim Ward of IBM's Olympia, WA office; Mark Cooper of the Safeco Insurance Company; and Steve Ballowe of SPSS, Inc.

AUTHOR'S ACKNOWLEDGEMENTS

Special thanks to Ann Blalock, Ernst Stromsdorfer and Gary Bodeutsch for intensive and thoughtful reviews of earlier drafts.

Carl Simpson

TABLE OF CONTENTS

LIST OF EXHIBITS vii

PREFACE ix

SECTION I. INTRODUCTION AND OVERVIEW

CHAPTER 1. REASONS FOR STUDYING GROSS IMPACT 1

Why Should Organizations Dedicated to the Delivery of Services Expend Resources to Collect and Analyze Data About Program Outcomes? 2

Why Extend Data Collection to Include a Post-Program Follow-Up? 3

Why Include an Employer Follow-up? 4

Research Questions Addressed by Gross Impact Analysis 4

The Relationship of this Guide to Post-Program Performance Requirements 6

CHAPTER 2. A QUICK MAP TO GROSS IMPACT ANALYSIS 7

Major Goals of the Gross Impact Model 8

Basic Ingredients in the Gross Impact Analysis of Participant Outcomes 10

Basic Ingredients in the Gross Impact Analysis of Employer Outcomes 11

State and Local Applications 12

What Types of Variables are Proposed for Measurement? 12

Specific Measures 15

CHAPTER 3. ISSUES UNDERLYING GROSS IMPACT ANALYSIS 19

PART I. DESCRIPTION OF PROGRAM OUTCOMES: USES AND LIMITS 20

Avoid Interpretations Implying Cause 20

Its All in How You Ask the Question 21

PART II WHAT MAKES VALID DIFFERENTIAL IMPACT ANALYSIS POSSIBLE? 23

Assessing Probable Causes Influencing Program Outcomes 23

What Major Sources of Bias Threaten Differential Impact Analysis? 24

How Can Bias from Non-Random Selection be Minimized? 26

Two Analysis Goals Confronting Potentially Severe Bias 28

PART III. DEMANDS ON THE NUMBER OF TREATMENT CONTEXTS AND ON THE SAMPLE SIZE OF EACH 29

Four Types of Measures 29

Demands on Sample Structure Made by Four Measurement Approaches 30

SECTION II. RESEARCH DESIGN

CHAPTER 4. RESEARCH DESIGN FOR A PARTICIPANT FOLLOW-UP 33

PART I. ALTERNATIVE LEVELS OF INVESTMENT IN THE ANALYSIS OF PARTICIPANT OUTCOMES 34

Three Levels of Investment in Gross Impact Analysis 34

PART II. IDENTIFYING THE POPULATION TO BE ANALYZED 36

Program Titles and Activities 36

Program Statuses 36

Duration of the Study 37

PART III. DESIGNING THE SAMPLE 39

Sample Size and Reliability 39

Level of Investment and Sample Size 40

When Should Samples be Stratified? 41

PART IV. SAMPLE SELECTION 43

Sampling Procedures 43

Responsive Rate and Sample Size 44

When are Participants Identified for Inclusion in the Sample? 44

PART V. WHAT FOLLOW-UP PERIOD IS RECOMMENDED? 45

The Basic Three Month Follow-up 46

Options for Extended Follow-up 46

Defining Termination 47

PART VI. DATA COLLECTION METHODS 48

Measuring Outcomes at Follow-Up 48

Measuring Program Variants 49

Measuring Control Variables 50

CHAPTER 5. RESEARCH DESIGN FOR AN EMPLOYER FOLLOW-UP 51

Alternative Approaches to the Analysis of Employer Outcomes 52

Identifying The Population to be Analyzed 53

Designing the Sample 53

What Follow-up Period is Recommended? 54

Data Collection Methods 56

CHAPTER 6. SUMMARIES OF RECOMMENDED EMPLOYER AND PARTICIPANT DESIGNS 59

Integration with Measurement of Post-Program Performance Standards 60

Profiles of Four Major Design Alternatives 61

- I. Minimal Investment, Termination Employer Follow-Up Descriptions 61
- II. Minimal State Oriented Participant Follow-Up 63
- III. Stepwise Employer/Participant Follow-Up 65
- IV. Local Emphasis Program Development Approach 67

Integrated Participant-Employer Designs 70

SECTION III. MEASUREMENT

CHAPTER 7. ISSUES UNDERLYING INTERPRETATIONS OF THE JTPA INTERVENTION 73
Conceptualizing the Nature and Location of JTPA Program Interventions 74
Program Impact Versus Effects of Individual Background Characteristics and Local Labor Market Characteristics 79
Are Outcomes at Termination and at Follow-up Produced in the Same Ways? 80
Issues Highlighted in JTPA Implementation Studies 81

CHAPTER 8. PARTICIPANT MEASURES 83
PART I. MEASURES OF PARTICIPANT OUTCOMES 84
Organizing A Limited Set of Potential Influences on Outcomes 86
PART II. MEASURING PROGRAM IMPLEMENTATION VARIANTS 87
Suggested Measures of Program Implementation Variants 88
PART III. MEASURING INDIVIDUAL TREATMENT VARIANTS 92
Measurement Sources 92
Suggested Measures of Individual Treatment Variants 83
PART IV. MEASURING CONTROL VARIABLES 101
Participant Characteristics 101
Labor Market Environment 106
Identification of Membership Categories 106
Recording Date Information 107

CHAPTER 9. EMPLOYER MEASURES 109
Separating Employer Outcomes from Outcomes for Others 110
PART I. THE NATURE OF EMPLOYER OUTCOMES 110
Measurement Strategy in a New Area of Study 111
Estimating Costs and Benefits Using the Gross Impact Approach 112
Tradeoffs among Different Costs and Benefits 112
Specifying Outcomes for Termination Employers 113
Outcomes for Participating Employers 114
Constraint, Investment, and Windfall among Participating Employers 115
Sources of Employer Cost and Benefit Measures 116
PART II. MEASURES OF EMPLOYER COSTS AND BENEFITS 117
Outcomes Which Apply to Termination and Participating Employers 117
Outcomes for Participating Employers Only 119

SECTION IV. ANALYSIS

CHAPTER 10. ANALYSIS PROCEDURES 123
Overview of General Approaches to Analysis 124
PART I. ANALYZING VARIOUS TYPES OF MEASURES 125
Levels of Measurement 125
Performing Multivariate Analysis with a Dichotomous Dependent Variable 127
Constructing and Testing "Dummy" Variables" 127
Constructing and Testing Interaction Terms 128
Reporting Standardized or Unstandardized Regression Coefficients 129
PART II. ESTIMATING CHANGE 130
PART III. ANALYZING A LARGE NUMBER OF POTENTIAL INFLUENCES ON OUTCOMES 131
PART IV. COMBINING DATA FROM DIFFERENT SOURCES 133

REFERENCES 135

SECTION V. APPENDICES

APPENDIX A. REVIEW OF SELECTED LITERATURE 141
Studies of Program Impact 142
Early JTPA Implementation Studies 143
Differential Impact Analysis 145
Broadening the Range of Outcome Measures 147
Employer Benefits 148

APPENDIX B. BACKGROUND TO GROSS IMPACT RESEARCH METHODS: SELECTED TOPICS 151
Sample Size and Sampling Error 152
Comparing the Three Major Types of Surveys 154
How Valid and Reliable are Surveys? 155
Differential Impact Analysis as a Statistical Model 158
Individual and Organizational Level Measures 159

APPENDIX C. SELECTED IMPLEMENTATION TOPICS 163
PART I. TOPICS CONCERNING THE ADMINISTRATION OF FOLLOW-UP SURVEYS 164
Locating Participants for Interview 164
Strategies Concerning Completion Rates 166
Introduction and Context Setting in Telephone Surveys 166
Who Speaks for the Employer? 167
PART II. ELEMENTS OF COST ACCOUNTING FOR IN-HOUSE FOLLOW-UP SURVEYS 168
Getting Started 169
Conducting Interviews 170
Fixed Costs 171

PART III. WHO SHOULD ADMINISTER DATA COLLECTION AND ANALYSIS? 173
Major Options for the Administration of Analysis Projects 173
Major Tasks to be Administered 173
Some Criteria for Deciding how to Admiinister Analysis Projects 174

APPENDIX D. ILLUSTRATIVE APPLICATIONS 177
Examples of Descriptive Applications 178
Examples of Differential Impact Analysis Applications 185

APPENDIX E. MEASUREMENT INSTRUMENTS 191
Criteria Under Which These Instruments Were Developed 192
Specific Instructions for Reading Follow-Up Interviews 193
SOME READY MADE DATA COLLECTION INSTRUMENTS 195
Participant 13 Week Follow-Up Survey Instrument 197
Employer Survey, Part 1. Outcome Questions for All Employers 205
Employer Survey, Part 2. Questions Describing the Firm 209
Employer Survey, Part 3. Questions for Participating Employers 212
Suggested Measures of Individual Treatment 216
Descriptions, Names, and Measurement Types For Measures Included in Appendix E 223

LIST OF EXHIBITS

Exhibit 2.1. A Summary of Some Definitions 9

Exhibit 2.2. Nature and Sources of Measures Which May Be Included in Gross Impact Analysis 14

Exhibit 2.3. Types of Variables Which May Be Included In Gross Impact Analysis 15

Exhibit 3.1. Defining Major Types of Bias In Differential Impact Research 25

Exhibit 3.2. Four Approaches to Measuring Program Variants, With Examples 29

Exhibit 6.1 Profiles of Four Major Design Alternatives 61

Exhibit 7.1 Examples of JTPA Interventions 75

Exhibit 8.1 Prioritized Participant Outcome Measures 84

Exhibit 8.2 Basic Elements of Differential Impact Equations 86

Exhibit 10.1 Suggested Statistics for Different Levels of Measurement 126

Exhibit 10.2. Original and Dummy Variable Forms of Marital Status 128

Exhibit B.1 Estimated Sampling Error 153

Exhibit B.2. Individual and Organizational Level Measures 160

(Exhibits D.1 through D.13, which report examples of gross impact analysis, are omitted from this List.)

PREFACE

This guide addresses a variety of design and measurement options. Many of these may profitably be combined or may be adopted separately. The major divisions in focus throughout the guide are:

- The measurement of *participant* and *employer* outcomes.
- *Descriptive* analysis of outcomes and multivariate *"differential impact analysis"* of the ways in which alternative forms of program implementation and treatment affect outcomes.
- Analysis uses and appropriate designs for *states* and for *SDAs.*

In addition, several other divisions occur within the guide. Optional designs are suggested for those deciding to invest larger or smaller amounts in measurement and analysis expenses. Also, although the major thrust of the guide is toward planning analysis design and measurement, some attention is paid to implementation issues. In particular, Appendix B addresses selected topics concerning the implementation and costs of conducting surveys, and Appendix C offers concrete examples of the questions and the findings from previous gross impact analysis.

The guide also includes ready-to-use interview instruments for participants and for employers. (See Appendix E.) It is hoped that these will subsidize states or SDAs who wish to analyze program outcomes but are hindered by the initial set-up costs.

The decision to address a variety of design and measurement alternatives directs the form in which this guide is written. Different chapters address employer and participant design and measurement. Different sections within many chapters address separate design options or other issues. These are labeled via extensive use of headings in hopes that readers will be able to locate the topics of greatest interest to them. The Table of Contents also includes some detail to help the reader locate issues. Chapter 2, A Quick Map to Gross Impact Analysis, offers an overview of the major components of the guide.

The various options discussed in this guide are unified by one overriding goal: that any analysis suggested here be oriented toward program development and improvement. The *forte* of gross impact analysis is the systematic description and comparison of the ways in which program implementation and treatment alternatives influence the success of the program in reaching desired outcomes. A major goal here is to support the development of program analysis which increases the knowledge base available to managers when they make program implementation and treatment decisions. Carefully designed program analysis, such as suggested here, can play a valuable role in looking to the future development of effective job training programs as well as in assessing how well programs have operated in the past.

SECTION I.

INTRODUCTION AND OVERVIEW

CHAPTER 1
REASONS FOR STUDYING GROSS IMPACT

Chapter 1. Reasons for Studying Gross Impact

This chapter discusses the types of questions that can be addressed with the gross impact approach, why service delivery organizations might wish to analyze those issues, and how the gross impact design coordinates with the measurement of post-program performance requirements.

WHY SHOULD ORGANIZATIONS DEDICATED PRIMARILY TO THE DELIVERY OF SERVICES EXPEND RESOURCES TO COLLECT AND ANALYZE DATA ABOUT PROGRAM OUTCOMES?

"Evaluation research" has often been viewed as remote from service delivery -- as serving distant purposes or as serving no purpose. However, the analysis of data on services and outcomes can be a valuable *management tool.* Perhaps this is why a recent survey finds many SDAs developing a special interest in systematic self-analysis (Seattle-King County, 1985). Examples of practical management uses of such analysis include:

- Continuously updated records on program outcomes provide a descriptive baseline against which to assess changes or stability in achieving intended program goals. This allows an SDA to trouble-shoot rapidly and it facilitates planning.

- When information on local program outcomes is collated at the state level, it provides a base to guide state disbursement of incentive and technical assistance funds, and can help determine whether states should make adjustments to SDAs' expected performance levels. The more carefully designed such comparisons are, the more valid and acceptable to SDAs these state decisions will be.

- *Systematic Descriptions* of program outcomes can focus policy planning discussions. Further, *analysis which compares the effects of program alternatives* can identify strong and weak areas of current services, in terms of their impacts on outcomes. The ability to focus change efforts on low performance areas and to identify high performance approaches as models for planning has the effect of continuously improving services. It amounts to "technological advance" for service organizations, where effective technology -- knowing what transformations produce desired outputs -- is difficult to develop with precision.

- The ability to analyze the effectiveness of program options carries with it the ability to conduct reliable evaluations of particular program innovations or pilot projects. This means that no separate evaluation contracts would be required to assess such projects.

- Collecting and analyzing data on program outcomes establishes the seriousness and goal oriented quality of service delivery organizations. Few techniques more effectively establish the seriousness and legitimacy of a manager than having a readily available data base which reflects program operations and can therefore be used to back up policy decisions.

- The ability to identify program strengths and weaknesses can also enhance staff morale. Information on how best to improve services provides staff with a sense of efficacy, the sense that they are able to affect the quality of their own work. Staff burnout has been identified as an ongoing problem under CETA (Franklin and Ripley, 1984). One partial solution is putting the tools for more effective management in the hands of local staff.

WHY EXTEND DATA COLLECTION TO INCLUDE A POST-PROGRAM FOLLOW-UP?

Even before the post-program performance standards became a likely federal requirement,[1] there was widespread interest in post-program follow-up data. The JTPA legislation points to the importance of post-program outcomes. Over 80% of states report either conducting follow-ups or making plans to begin them (National Governor's Association, 1985). In addition, SDAs' focus on service delivery and states' focus on measuring and improving SDA performance can be effectively supported through analysis of follow-up data.

With the advent of post-program performance requirements, collecting extended follow-up data becomes marginally efficient. Additional measures can be integrated with the required follow-up. Even without this incentive, however, there is reason to entertain post-program measures.

Follow-up provides valuable information not otherwise available. Employment continuity and intensity are important program outcomes, measurable only after a period following termination. Similarly, only after a participant has experienced a position for a time can he or she report whether the job is using skills learned from skill training, or how likely he or she is to remain in the job. It is also only after a follow-up period that employers can report how likely the participant is to be retained or promoted, how well trained the participant was, or how satisfied the employer is with the quality of the participant's work.

Follow-up data also make comparisons among service providers more reliable. Termination can be a slippery concept, susceptible to multiple definitions and to some degree of manipulation by program operators, undermining comparability. This problem is avoided with follow-up data.

Finally, and most important, viewing outcomes at some distance after the program intervention helps illuminate the most important program outcomes -- the lasting change which JTPA wishes to create. Follow-up measures have been shown to be more accurate than termination measures as *indicators* of long term net improvement in employability and income (Gay and Borus, 1980; Zornitsky, et al., 1985b).

Being *hired* at program termination and being *retained* are somewhat different matters. For example, training in job interview skills may increase the likelihood of finding a job rapidly, but not increase retention. Skill training may have the reverse effect. Further, different program activities tend to produce somewhat different employment trends. Employer-based activities may produce higher employment rates at termination than at follow-up, whereas classroom training is likely to produce the reverse pattern. Indeed, a follow-up of terminations that did not result in immediate job placement can prevent premature rejection of a particular program because of a temporary hiring dip in a particular occupation or industry.

---

[1] At this writing, post-program performance requirements are likely but not definite. The design and measurement recommendations in this guide are consistent with probable Department of Labor requirements. That coordination is based on *draft documents* of proposed JASR requirements and of a post-program data collection supplement made available by the Department of Labor. The remainder of this guide is written on the assumption that post-program data collection will be required and that these requirements will take the form proposed at the time of this writing, January, 1986.

These considerations mean that follow-up data place managers in a better position to assess the relative effectiveness of JTPA program variants.

WHY INCLUDE AN EMPLOYER FOLLOW-UP?

The new emphasis on the private sector dictates an interest in potential benefits to employers. Employers can report their perceptions of how well trained JTPA hires are, whether turnover is reduced or increased by hiring JTPA participants, how useful JTPA screening is, and the like. They can also indicate their satisfaction with particular JTPA services or the types of services they feel they or their JTPA hires need.

In cases where employers participate in the JTPA program by hiring a participant into on-the-job training or work experience, costs as well as benefits become important. Employers receive payments subsidizing these hires. In employers' perceptions, how great are the offsetting costs of hiring JTPA participants rather than other workers?

Employers are also in the unique position to provide information on certain valued participant outcomes; in particular, the likelihood that the participant-employee will be retained or promoted and the likelihood that layoffs at the point of follow-up are permanent or temporary. Employers who participate in providing services can also report the nature of those services -- in particular the amount and types of training provided.

JTPA programs may be well advised to conduct systematic employer follow-ups for strictly programatic reasons in addition to their use as measures of program outcomes. Employer follow-ups can include a brief retrospective discussion of each JTPA placement. This allows the employer to discuss problems that might hinder future participation and to consolidate or expand future plans with JTPA. Similarly, such a discussion can be part of the quality review undertaken before the service providers decide whether to encourage future participation by a particular employer.

RESEARCH QUESTIONS ADDRESSED BY GROSS IMPACT ANALYSIS

In general, the rationales given above for analyzing post-program outcomes for participants and employers refer to goals which can be reached using the gross impact analysis approach presented in this guide. It is also important at the outset of the guide to explicate what types of analysis questions can and cannot be addressed using this approach.

Questions Gross Impact Analysis Cannot Address

Among the wide range of *possible* impacts of JTPA, only a smaller set can be addressed using the gross impact approach. Several very important society wide goals of job training programs are essentially impossible to study *definitively*, because legislation which improves the situation of some individuals may be creating or overlooking problems among other individuals. These impacts include increasing national productivity, reducing total national unemployment, reducing average job turnover time, and improving the skill level and therefore the flexibility of the overall labor force.

In addition, gross impact analysis cannot draw any conclusions concerning the types or degree of change *caused by participation* in the JTPA program. This question can only be addressed by net impact studies, which compare program participants with similar individuals who did not participate in JTPA programs (Johnson, 1986). This distinction is so important that it is the basis on which the net and gross impact analysis guides were named. Gross outcomes describe total post-program outcomes; net impacts estimate the proportion of total outcomes which may be uniquely attributed to participation in a JTPA program.

This means that gross impact studies cannot estimate the extent to which participation in JTPA changes individuals, the cost-effectiveness of JTPA, the time it takes for participants to repay the cost of the program in taxes generated by program success, or the impact of the program on welfare roles. Nor can gross impact analysis ask whether some category of clients is more benefited by JTPA participation than some other category of clients. These are all net impact questions concerning unique effects of the overall JTPA program; they require comparison with non-JTPA-treated individuals.

Finally, gross impact analysis can add little to the diagnosis of program implementation for a single service delivery organization. That is the domain of process analysis. Post-program outcomes can be described for one organization, but assessing how the organization operates requires either comparison among many organizations (the gross impact approach) or in-depth analysis such as described by the process analysis guide (Grembowski, 1986).

Questions Gross Impact Analysis Can Address

There are two broad categories of analysis goals for which the gross impact approach is well suited: 1. the *description of a broad range of program outcomes* , and 2. estimating the unique impact on outcomes produced by *alternative methods of delivering services.*

Description of Post-Program Gross Outcomes does not infer causation. This guide will refer to the description of *gross outcomes,* avoiding the word impact as a reminder that no cause-and-effect impact can be identified from descriptive analysis.

However, gross impact analysis is well suited to describing a wide range of employer and participant outcomes, with results available in a relatively short time. These results can be evaluated against manager's goals. JTPA performance standards are one type of descriptive gross outcome for which precise goals have been established. For other outcomes, descriptive data may help establish reasonable baseline expectations. Gross outcomes can also be used as tools to identify problem areas. For example, employers' evaluations of various aspects of JTPA are descriptive outcomes. They can also show changes over time in program outcomes. And the ability to measure relatively numerous and detailed outcomes provides a way to describe the range of program outcomes and the ways in which programs are achieving their impacts. For example, outcomes can include whether employment is training-related, the quality of jobs such as fringe benefits and likelihood of promotion, and employer willingness to participate in future JTPA programs.

Differential Impact Analysis is a method for rigorously comparing alternative approaches to delivering services. Different program treatments can be compared to assess whether, other things equal, *one or more alternative is more effective than others in producing desired outcomes.* That is, the unique impact of each *program variant* can be estimated in comparison to all other program variants in practice during the analysis. Participants experiencing each program variant act as a comparison group for those experiencing other variants. This opens the way to a wide range of analysis questions which might be asked by managers. The specific questions depend on what program variations exist and which areas of service managers are most interested in developing.

Within the constraints of sample size, these same questions can be asked for *particular populations of participants:* which treatment modes are most effective for target group A, and which for group B?

Comparisons can also be made among *service providers.* This means that states can improve the reliability and meaningfulness of comparisons made among SDAs, and can also identify especially valuable directions for program technical assistance efforts, as long as the influences of labor market conditions and other important differences among SDAs are taken into account.

THE RELATIONSHIP OF THIS GUIDE TO POST-PROGRAM PERFORMANCE REQUIREMENTS

When this guide is published, it is very likely that the U.S. Department of Labor (DOL) will have required SDAs to measure the post-program employment and earnings of former JTPA participants.

The measures selected as post-program performance standards are the most important of all gross outcomes. They consequently serve as the most valuable dependent variables for differential impact analysis. That is, these are the post-program outcomes managers will most wish to maximize when developing and refining service delivery policies.

The requirement of interviewing a sample of participants from each SDA at follow-up means committing considerable resources. That in turn means that in many cases, for states or SDAs to conduct entirely separate analysis of program outcomes becomes financially prohibitive. However, when gross impact analysis is conducted in conjunction with required post-program follow-up, many costs are shared, making gross impact analysis less expensive than it would be without the DOL requirement.

Therefore, this guide has been structured so as to assure consistency with DOL requirements. Designs for employer analysis cannot be integrated with performance requirements, which demand participant contact only. However, where possible, the research design and measurement recommended here can be integrated with measurement to satisfy DOL requirements.

CHAPTER 2

A QUICK MAP OF GROSS IMPACT ANALYSIS

Chapter 2. A Quick Map of Gross Impact Analysis

This chapter lays out the basic terrain covered by the gross impact model. It defines the terms used in this volume and identifies the basic ingredients of the gross impact approach. The research design elements and types of measures recommended are also summarized.

The chapter is cryptic; topics are covered briefly, with little explanation and without reference to previous literature. Each design and measurement topic summarized here is treated in greater detail later in the volume.

MAJOR GOALS OF THE GROSS IMPACT MODEL

- Rapidly developing reliable **descriptive information on participant outcomes at follow-up**, a period after termination. That information includes a variety of employment and earnings outcomes for all participants or for particular subgroups of participants. This guide also includes optional measures of a *wider range of participant outcomes.*
- Assessing the **outcomes of the program for employers** in their roles as service providers, in on-the-job training and work experience; as employers of job-ready participants who received only job search assistance; and as employers who hired classroom training (CT) participants or others at their program termination.
- Analyzing the **differential impact of variant forms of program implementation and of the services provided.** The program variants which can be compared include basic divisions such as *program activities* or *service providers;* the *policies or practices* that service providers adopt to implement intake, referral, treatment, and placement; and specific variations in the *treatment each participant receives.*

Who Can Use Gross Impact Analysis

Gross impact analysis has value at all levels of the JTPA delivery system. However, it is pitched primarily to the local program management level. Its major thrust is to develop information which managers can use to plan more effectively. In particular, it provides access to well measured program outcomes and provides a method to assess how effectively the available service delivery options produce desired program outcomes. Such differential impact analysis is sometimes referred to as "fine-tuning" the organization (Rossi and Freeman, 1982).

EXHIBIT 2.1. A SUMMARY OF SOME DEFINITIONS

Gross Impact Analysis Model

An analysis approach for gathering original survey data from participants and/or employers, measuring a range of post-program outcomes and selected aspects of program implementation and treatment, and analyzing that information within appropriate limits in the absence of an untreated comparison group. Analysis includes two levels of complexity: description of gross outcome and differential impact analysis.

Descriptive Analysis of Gross Outcomes

Reports of the nature and level of participants' post-program labor market experiences and of employers' experiences with JTPA and JTPA participants, without comparisons or causal interpretation.

Differential Impact Analysis

Comparisons of the relative effectiveness with which different forms of program implementation and services options produce desired program outcomes, using multivariate statistical techniques to adjust for effects of the labor market environment and the background characteristics of individual participants.

Program Variants

The term *program variants* refers to the alternative forms of agency implementation or services provided to each individual participant that can be analyzed using the differential impact analysis approach. These are the explanatory policy variables in differential impact analysis. They may be policies directing service delivery, typical service provider practices, the basic program activities to which individuals are assigned, or variant forms of those activities.

Some Program Components:

- The term *program*_ refers to the set of policies and practices implemented under JTPA.
- *Program Implementation* refers to the organizational structures, patterns, and policies through which JTPA services are delivered.
- *Program activity* refers to the basic JTPA treatment modalities, on-the-job training, classroom training, etc.
- *Services* are inputs provided to assist a recipient, such as financial resources or leverage, advisement, referral, and training.
- *Treatment* or *intervention* means the set of services rendered to one recipient by JTPA service providers.
- A *service provider* is an agency, SDA or subcontractor, charged with direct delivery or procurement of services.
- A *participating employer* is one who receives funds from JTPA in exchange for providing training, experience, or employment.
- A *termination employer* is one with whom a participant is employed at termination, regardless of whether that employer also participated in delivering services.

Abbreviations Used In This Guide

OJT	On-the-job training	*JSA*	Job Search Assistance
WEX	Work Experience	*CT*	Classroom (skill) training

This approach can also assist state JTPA policy makers in their oversight role. It provides a tool for judging relative SDA performance levels more reliably than by using descriptive status reports alone, increasing the validity of adjustments which might be made to performance standards. And it can identify effective program practices which might be recommended as part of technical assistance efforts.

BASIC INGREDIENTS IN THE GROSS IMPACT ANALYSIS OF PARTICIPANT OUTCOMES

What Overall Designs Are Recommended?

All, or a sample of all, participants terminating in any status from the program activities which state or SDA users decide to analyze would be studied. Three major options, varying in the level of investment required to conduct the analysis, are:

1. Descriptive analysis only, including a wide range of outcome measures.

2. Differential impact analysis comparing service provider implementation measures. This is most appropriate for state level analysis.

3. Differential impact analysis including detailed individual treatment measures. This design is especially well suited to intensive local analysis.

What Would be Measured?

Participant outcomes are measured in all cases. Differential impact analysis also requires measurement of the factors which might affect outcomes, including program variants and control variables measured to guard against bias. (See Exhibit 2.2.)

How Would Participant Outcomes Be Measured?

Telephone interviews would be used to gather participants' reports of their labor market experiences after JTPA.

How Would Program Variants Be Measured?

Program Variants include several types of measures and several data sources which a state or SDA *may* choose to include. The major types of measures are implementation variants and individual participant services.

Program Implementation Variants are measured for each service provider. Each organization receives one score on each measure. Such measurement may occur once only, with occasional updates when forms of implementation change.

Participant Services are measured from the standpoint of the individual, but may come from three different sources, listed below. Information from each source is attached to the individual data record, becoming a characteristic of the services that individual received.

1. Services such as basic program activities and support services are recorded in an agency MIS.

2. Characteristics of the treatment (e.g. the types of training an OJT employer uses or the classroom trainer's instructional approach) may be measured using one-time-only or annual interviews with the employers or trainers.

3. Characteristics of the treatment can also be gathered from participant reports, preferably at termination, when experiences are fresh, but optionally during follow-up interviews.

Statistical Control Variables may come from MIS files in the case of participants' demographic backgrounds, from service providers in the case of intake selectivity measures, and from state agencies or census files in the case of the labor market and demographic environment.

How Long Would the Post-Program Follow-Up Period Be?

A three month follow-up is recommended as the most basic option. That period is consistent with DOL requirments for measuring post-program performance. It is also long enough to avoid most of the immediate post- program employment instability, yet brief enough to minimize difficulties in locating clients. Recommended methods for longer follow-ups are also discussed.

How Would the Data Be Stored?

The answer to this question depends on the data processing systems used by each state or SDA. Ideally, these data would be integrated with MIS files in a continuous-update system, but many options are workable.

How Would the Data Be Analyzed?

For descriptive uses, percentage distributions and averages are most informative. For differential impact analyses, data from different sources would be merged into individual participant files, and multivariate analysis performed.

BASIC INGREDIENTS IN THE GROSS IMPACT ANALYSIS OF EMPLOYER OUTCOMES

What Overall Designs Are Recommended?

Three possibilities are outlined, depending on the uses intended for these data.

1. A **survey of termination employers** who hired participants from any program activities would measure employer outcomes revolving around the work performance of the former participants who were hired by each employer. Employer data would not be integrated with participant data.

2. A **survey of participating employers** who delivered OJT, WEX, or other services would include employers in the program activity being analyzed, regardless of the termination status of the participants placed with them. Employer benefits and costs of participation would be analyzed.

3. The third possibility is a **multi-purpose employer survey**, combining employer and participant data. In this case, employer outcomes are measured and employers are also asked to describe the training or other services they provided to participants. These may be treated as employer costs of participation and as program variants to be included in differential impact analysis of participant outcomes.

How Would Employer Outcomes Be Measured?

Follow-up interviews should typically be conducted over the telephone, although possible exceptions are discussed in Chapter 5.

How Long Would the Employer Follow-Up Period Be?

A three month follow-up keeps employer and participant data as nearly comparable as possible. However, interviews with participating employers who did not retain their participant could be conducted shortly after the contract ends. Also, the sequential combination of employer and participant interviews is extremely efficient, but cannot be integrated with DOL performance requirements at the three month follow-up. Therefore, states or SDAs wishing to perform longer follow-ups may wish to include employer at that point.

How Would the Data be Stored and Analyzed?

Employer data can be analyzed independently. They can also be attached to the data files of individual participants for a combined analysis.

STATE AND LOCAL APPLICATIONS

Most analysis goals and research design elements are identical for states and SDAs. The greatest strength of the differential impact approach is its ability to identify ways in which direct delivery of services can be improved. Both local service providers and state technical assistance providers have an interest in such information.

However, some differences exist in the designs which are most useful to state or to SDA level analysis. State or multi-SDA analysis projects are able to compare post-program outcomes of SDAs, adjusting for other influences such as local labor market conditions and participant characteristics. Further, only designs including a considerable number of service providers are able to perform differential impact analysis of service provider implementation variants. (See Chapter 3.) Therefore, one state-wide design recommended as highly efficient would perform differential impact analysis on a limited range of outcomes using a survey of service providers to measure implementation variants for analysis.

SDA level analysis can compare subcontractor implementation forms if a large enough number exist. In addition, local efforts are in a better position to collect detailed information on services to individual participants. This makes in-depth differential impact analysis of particular program activities an appealing local option.

WHAT TYPES OF VARIABLES ARE PROPOSED FOR MEASUREMENT?

Any analysis project can be successful only if the correct factors are measured and the measures are correctly constructed. This guide includes discussions of issues underlying the selection of measures, recommendations for more specific factors which might be measured, and ready-made survey instruments which can be used to collect data on many of the factors discussed as possible measures. (See Appendix E.)

Exhibit 2.2 gives an overview of the various types of measures which may be involved in gross impact designs. Measures are grouped according to the purpose each measure serves during analysis (outcome, program variant, control variable), the aspect of the JTPA program being measured, and the source of each measure. The most critical division is by purpose, since the questions to be addressed in studying gross outcomes require the definition and measurement of outcomes, possible influences on those outcomes, and control variables to protect against bias.

Any given differential impact analysis involves one outcome variable and some number of program variants, along with the control variables included to protect against selection bias or other sources of error. Within this structure, each test variable represents an hypothesized effect on the outcome in question.

Although the array of different measurement sources listed in Exhibit 2.2 may seem complex, the diversity is helpful in two ways. First, each factor is measured using the most reliable source. Second, when reliable options exist, variables are listed under the least expensive source. Surveys of service providers are inexpensive because their numbers are small compared to participants. Participant treatment records kept by service providers are both more reliable and less expensive than measures included during follow-up interviews of former participants. Participating employers can be interviewed when they first agree to participate, or at follow-up. Once- only interviews are more efficient for employers who enroll multiple participants.

Variables measured as *controls against bias* can come from many sources, depending on the specific design of the analysis. Employers, service providers, and participants all have something to report on the selection process. Service providers also develop policies which affect selection bias. MIS files and published demographic and labor market data report standard variables known to affect labor market success and therefore necessary to include in differential impact analyses. One strength of the gross impact approach is its *flexibility to measure multiple indicators of selection.*

EXHIBIT 2.2. NATURE AND SOURCES OF MEASURES WHICH MAY BE INCLUDED IN GROSS IMPACT ANALYSIS

	The Purpose of Each Measure and Its Relation to JTPA				
	Outcomes		Program Variants		
Source	Participant	Employer	Implementation	Services	Controls against bias
Survey of service providers			E		I
Participant treatment kept by service providers		C	F*	G	J
Standard MIS files					K
Participant follow-up surveys	A			**	
Participating employers				H	
Employer follow-up surveys	B	D			L
Published data by locality					M

Note: The lettering above provides a reference organizing measures in exhibit 2.3

* Individual treatments may be aggregated to indicate typical agency patterns.

** Selected service variants can also be measured through participant surveys.

SPECIFIC MEASURES

One goal of this guide is to open analysis options and stimulate thinking in a number of areas. This means that many specific measures are suggested. These are presented in Chapters 8 and 9, and Appendix E, and are not repeated here. However, Exhibit 2.3 offers an overview of the types of measures suggested. These are listed under the broad categories defined in Exhibit 2.2.

EXHIBIT 2.3. TYPES OF VARIABLES WHICH MAY BE INCLUDED IN GROSS IMPACT ANALYSIS

A. PARTICIPANT OUTCOMES MEASURED IN PARTICIPANT FOLLOW-UP SURVEYS

(Numbered below in order of recommended priority.)

1. Required post-program performance standards.
2. Other core measures explicit in the JTPA mandate, measured pre- and post-program, to allow calculation of change.
    - Employment.
    - Earnings.
    - Welfare dependence.
3. Measures of skill transfer and utilization.
4. Measures of job quality, indicating primary versus secondary labor market jobs.
5. Measures characterizing those not employed or not retaining jobs held at termination.
6. Subjective orientations of participants

B. PARTICIPANT OUTCOMES MEASURED IN EMPLOYER FOLLOW-UP SURVEYS

- Retention of placement employment.
- If retained: job qualities.
- If not retained: why not?

C. EMPLOYER OUTCOMES VIA PARTICIPANT TREATMENT RECORDS

- For participating employers only: direct monetary benefits.
- Amount of applicant screening provided by the agency.

Exhibit 2.3 (continued)

D. EMPLOYER OUTCOMES MEASURED IN EMPLOYER FOLLOW-UP SURVEYS

For All Employers:

- Participant performance on the job.
    - -- Skill level.
    - -- Job performance.
    - -- Supervision ease and work habits.
    - -- Personal adjustment on the job.
- Job retention and reasons for non-retention.

For Participating Employers Only:

- Perceived benefits and costs of participating in JTPA.
- Perceived value of JTPA services.
- Reported costs involved in providing services.

E & F. PROGRAM IMPLEMENTATION VARIANTS

(Direct characterizations of implementation (type E) are measured by interviews with service providers. Aggregated measures (type F) are calculated from participant treatment records.)

- Basic organizational composition.
- Service delivery framework.
    - -- Intake policies and practices.
    - -- Quality control over referral and program activity mix.
    - -- Policy toward ancillary support services.
    - -- Exit practices.

G. PROGRAM SERVICES VARIANTS VIA PARTICIPANT SERVICE RECORDS

(Alternatively, some service variants may be measured during interviews.)

- Screening, selection, and intake services.
- Referral to basic program activities.
- Treatment intensity and completion.
- Trainer characteristics.
- Ancillary support services.
- Individual treatment cost.
- Program exit and job search.

Exhibit 2.3 (continued)

H. PROGRAM SERVICES VARIANTS VIA PARTICIPATING EMPLOYER SURVEYS

- Nature of training and placement job.

I. CONTROLS AGAINST BIAS VIA SURVEY OF SERVICE PROVIDERS

- Selection policies and practices.

J. CONTROLS AGAINST BIAS VIA PARTICIPANT TREATMENT RECORDS

(If necessary, participant follow-ups can estimate many of these.)

- Diagnosis of pre-program need for intake and employment barriers.
- Participant self-selection into program alternatives.
- Individual route into JTPA and particular JTPA activity.

K. CONTROLS AGAINST BIAS VIA STANDARD MIS FILES

- Participant background characteristics.

L. CONTROLS AGAINST BIAS VIA EMPLOYER FOLLOW-UP SURVEYS

- Employer influence over the hiring and referral process.

M. CONTROLS AGAINST BIAS VIA PUBLISHED RECORDS FOR LOCALITIES

- Local labor market conditions.

CHAPTER 3

ISSUES UNDERLYING GROSS IMPACT ANALYSIS

Chapter 3. Issues Underlying Gross Impact Analysis

This chapter discusses the basic ideas underlying *descriptive analysis* of program outcomes and *differential impact analysis.* Each approach is discussed both in terms of its potential uses and in terms of factors which can limit its validity. The chapter ends with a discussion of a major issue dictating state and local settings in which use of differential impact analysis is appropriate: the number of discrete service providing agencies required to analyze variables of different types.

PART I.

DESCRIPTION OF PROGRAM OUTCOMES: USES AND LIMITS

Descriptive analysis takes its name from its goal of examining outcomes without making causal attributions. Descriptive patterns may be reported on the basis of data covering all participants or may be estimated from a sample of participants. Where sampling is involved, proper procedure will generate unbiased estimates of patterns characterizing all participants and information on how accurate those estimates are.

Descriptive data are relatively easy to collect and report, but also easy to misinterpret. This discussion therefore takes two directions: identifying ways to make gross impact description most useful to JTPA managers, and identifying the major limits on its valid interpretation.

AVOID INTERPRETATIONS IMPLYING CAUSE

As the term *description* indicates, the primary limitation on descriptive data analysis is that it involves none of the research design or analysis techniques for explaining causal relationships. The major reason for this limitation is that descriptive analysis offers no comparisons. For example, if we learn that employers are highly satisfied, we cannot know whether the reason is the reimbursement, the friendly service, the qualities of participants, cognitive dissonance, or a general tendency to answer positively. We can guess, but the research findings offer no guidance until comparisons are made—in this example, comparisons between employers training more and less qualified participants, with higher and lower reimbursement levels, etc.

This limitation does not mean that managers must refrain from interpretation. We all interpret the world daily. It means that managers *must not assume that the findings imply a particular interpretation.* Thus, the first of the following two statements a JTPA manager might make is flatly incorrect, while the second could be correct:

1. "We find employers to be highly satisfied with JTPA, demonstrating that we are sending them the types of employees they want."

2. "We find employers to be highly satisfied with JTPA. In my opinion, this is true because we are sending them the types of employees they want."

Statement 2. avoids incorrect causal attributions while also stating a possible interpretation which could be examined through further analysis. The value of the descriptive finding is that it *identifies the facts the manager may work with and may attempt to explain.* We can learn how satisfied employers are. The limitation is that the findings do not themselves offer any causal explanation for the level observed.

The most common error in interpreting descriptive data on job training programs is to assume that outcomes described following the program are caused by the program. In the heat of political battle, I may say "Look what our program has accomplished; we have 84% placement rates!" In so stating, I may be taking credit for upswings in the economy, for individuals who recovered from temporary unemployment, and for random change, as well as for cases where employment was produced by the program. Similarly, if I claim that one service provider is "better" (causes greater success) than another on the basis of descriptive findings, I err by assuming that the difference was produced by program services alone, which cannot be demonstrated using descriptive statistics.

ITS ALL IN HOW YOU ASK THE QUESTION

The first approach to both the limits and the potentials of descriptive analysis is to ask questions which are meaningful without demanding more complex comparisons than allowed. Some questions involve no interpretation; they simply seek *baseline descriptive information.* Other questions may be worded specifically enough that a descriptive answer will assist the analyst in *developing or confirming explanations.* The following types of questions illustrate.

Does It Appear That Program Goals Are Being Met?

If I know roughly what *levels of program outcomes* are expected, measuring outcomes lets me know whether I am in condition red, yellow, or green. Descriptive levels do not tell me why outcomes are higher or lower than expected, or whether my program itself has much to do with producing those outcomes. However, they tell me whether I need to look for factors creating low outcomes, whether high or low outcome levels are concentrated in particular program activities, whether my organization is in better shape with regard to some outcomes than others, and the like. That is, descriptions of outcome levels can let managers know whether to worry, and which program areas to worry about most actively.

For some post-program outcomes, those mandated as post-program performance standards, clear expectations will be established. When expectations are unclear, descriptive measures can help establish reasonable state or local baseline expectations. These would constitute first approximations which might be improved upon in subsequent efforts.

Does Any Service Provider Appear Worth Learning More About?

One particularly useful application of descriptive analysis as a first approximation is the *comparison of SDAs or subcontracting service providers.* Such comparisons should be made with great care, since agency performance levels are influenced by factors over which program operators have no control, such as the local economy, or may result from policies not intended by the act , such as increasing performance rates by serving those with least need. Descriptive differences point out where further investigation might be most useful, helping *pose* questions correctly rather than answering them.

Is There Any Apparent Change Over Time?

Descriptive outcome figures kept over time, each month for example, can be used to form a baseline series indicating stability or change in services provided and program outcomes. Such a "time series" can sometimes alert managers to unexpected changes. It can also provide a relatively inexpensive first approximation of the effects of major program changes made during the time series.

Investigation of Specific Propositions

One major strategy of multivariate analysis is to test a particular interpretation by seeing whether competing explanations for the observed findings can be eliminated. This tactic is not available for descriptive analysis. However, the same general strategy may be followed by posing questions thoughtfully and specifically enough to reduce the range of findings which would be consistent with the particular explanation proposed.

There is little value in asking broad questions such as "Does OJT produce more placements than CT?" during descriptive analysis. Too many different interpretations could reasonably explain either positive or negative findings. However, specific propositions direct expectations to only a few findings. *If* the expected finding occurs, we have greater faith in the correctness of the proposition guiding the analysis.

For example, if I identify some JTPA program activities as skill training programs, I will expect that a disproportionate number of post- program job placements will be in the skill area. I have no *a priori* way to set expected levels, but descriptive findings are nevertheless interpretable. If only 2% of workers in my area are cashiers, and only 6% of my CT participants have previous experience as cashiers, then a finding that 65% of employed graduates from my cashier training class are cashiers suggests that the program is working in the way I envisioned. This does not indicate how well the program works; only that my proposed explanation about the way it works is supported.

Another example involves the question: what accounts for non-retention of jobs held at JTPA termination? Several specific propositions are easy to imagine, each suggesting its own specific measures. For example, if JTPA participants lack the ability to learn complex skills, instances of non- retention should occur most often when the training or the job involved complex skills or where the participant's pre-program skills were weakest, and employers should often report that the participant was unable to perform complex tasks. If these variables are measured along with others indicating alternative explanations, managers can assess which explanations account best for the patterns observed.

As a final example, one may argue that JTPA should move participants into primary labor market positions (Taggart, 1981). One could examine the degree to which this occurs by measuring qualities of post-program jobs which define the primary labor market, including benefits packages, job security, presence of a promotion ladder, etc. (see Doeringer and Piore, 1971; Vermeulen and Hudson-Wilson, 1981). Findings would not indicate the degree to which JTPA treatment caused the job quality mix observed, but they would recommend greater or lesser concern about program quality depending on the number of jobs exhibiting the desired qualities.

Perceptions Held by Employers and Participants

Some questions are inherently descriptive. If I wonder what importance employers place on various qualities of individuals they hire, I can ask them to tell me. Although data on such perceptions may be limited by incorrect self-knowledge or by misleading responses, these perceptions are appropriately interpreted in their descriptive form. The same is true of participants' job satisfaction or other participant perceptions in which JTPA managers may have interest. Similarly, employers' satisfaction with JTPA and their perceptions of the costs and benefits of participating in OJT or WEX may be taken at face value, as long as one recognizes that the information indicates no more than perception, and that perceptions do not necessarily reflect program impact.

PART II.

WHAT MAKES VALID DIFFERENTIAL IMPACT ANALYSIS POSSIBLE?

ASSESSING PROBABLE CAUSES INFLUENCING PROGRAM OUTCOMES

The goals of differential impact analysis include reliably describing differences in post-program outcomes and also identifying the *probable causes* of those differences. In non-experimental research, identifying causal relationships is problematic. However, quasi-experimental research designs such as those recommended by Campbell and Stanley (1966), can considerably increase our confidence that we have reliably identified the major causes of differences we observe. (See also Cook and Campbell, 1979; Caporas and Roos, 1973).

Identifying probable causal connections is valuable to program managers because changing a factor that has causal influence on program outcomes is likely to change the level of those outcomes. Each of the steps involved in differential impact analysis has the goal of increasing our confidence that we have identified those program variants which do have a direct influence on program outcomes and are therefore useful to program managers in improving their programs.

This guide can only summarize some major characteristics of non-experimental research designed to increase the analyst's ability to identify causal relationships. There are also useful references available on this complex subject (Blalock, 1964; 1971). The statistical model underlying differential impact analysis is discussed briefly in Appendix B.

Research into causal relationships begins with comparisons. To determine whether program option A is better than option B, one must identify a criterion of comparison (e.g., job retention) and compare options A and B on that dimension. Options could be basic program activities, or optional variants of the same activities. These comparisons should be selected so that a causal interpretation is reasonable. This is where past research findings, economic theory, and managers' knowledge of programs come into play. If answering a question in causal terms would fly in the face of logic or of established information, the question probably should not be posed as part of a differential impact analysis.

In addition, to convincingly establish that a relationship is causal, findings from our comparisons must hold up after competing explanations have been eliminated. Each time we identify a plausible competing explanation, test it, and find that it does not explain away the difference between options A and B, we increase our confidence in the causal association between program variant A/B and the outcome in question. The goal of quasi-experimental research is to eliminate all important measurable alternative explanations. That goal is never reached, but we can eliminate *many* important alternative explanations. These include both factors of interest to the analyst, such as other *program variants confounded with the one being tested,* and *"control" variables* known to affect the outcome in question, such as age or gender.

Classical experiments attempt to eliminate competing explanations by controlling variants other than the A/B comparison of interest and by randomly assigning individuals to variants A and B, hoping to produce groups equivalent in all regards except the variant under study.[1] Quasi-experimental research occurs in settings which allow neither the control of variants other than those directly under study nor the random assignment of participants to program variants. Instead, multivariate statistical techniques are used to determine whether alternative explanations are able to undermine our confidence in findings. The most widely available multivariate technique, ordinary least squares multiple regression, is adequate for most differential impact analysis.

The primary strategy of differential impact analysis is to *utilize each program variant as a comparison group for each other variant.* Except where program variants are too highly correlated with each other or with participant background characteristics, multivariate analysis can estimate the unique effects of each.

WHAT MAJOR SOURCES OF BIAS THREATEN DIFFERENTIAL IMPACT ANALYSIS?

The comparisons demanded by the analysis goals and the method of collecting data, in this case surveys, determine the major threats to the validity of differential impact analysis. These are summarized in Exhibit 3.1. These are each sources of *bias*, as opposed to *random error.*

The term *bias* refers to error which consistently misdirects research results in a particular direction. Like a compass with a metal object nearby, readings from the analysis are distorted in a consistent direction. To correct the findings, one must remove the object or adjust for its influence. Random error differs from bias in that it takes no particular direction. Random error can be as serious as bias if it is large. However, techniques for minimizing random error are well developed in survey research: i.e., careful measurement techniques and properly constituted samples.

Each type of bias listed in Exhibit 3.1 can distort estimates of program outcomes. Three of these types, censored samples, non-response bias, and response bias, can be prevented or reduced during the process of sample definition and data collection. Selection bias and bias from confounded program variants are combatted during multivariate analysis. This means that descriptive analysis, which does not employ multivariate techniques, is always subject to serious bias. Differential impact analysis is able to reduce, but not eliminate, these biases during analysis.

Bias is reduced when equations include measures indicating selectivity, and measures indicating pre- and post-program conditions. However, many selection biases are unknown or cannot be measured, making statistical adjustments difficult. Therefore, selection bias is the most serious analytic problem, as well as the most difficult to diagnose.

---

[1] Most of the questions addressed using differential impact analysis (comparing effects of treatment variants) cannot realistically be studied using the experimental design. There are four reasons:
a. Political and ethical questions prevent the frequent or casual use of random assignment.
b. Any one experiment is capable of assigning individuals to a limited number of treatment options. Thus, a large number of separate experiments would be required to examine the range of treatment options differential impact analysis can examine somewhat less reliably.
c. In many cases, random assignment is a programatically meaningless approach, and as such lacks validity. (Findings could not realistically be generalized to non-experimental situations.) If a client clearly needs or requests a certain type of services, random assignment to other services would constitute inadequate program response. The results of such an experimental approach would be safely generalizable only to programs which place participants without properly diagnosing their needs.
d. Random assignment to treatment variants would in some cases cause uncontrollable selection bias on the part of participants, thereby eliminating the advantage of random assignment. In particular, OJT referrals made by employers will enroll if they are randomly assigned to OJT but will tend not to enroll if they are randomly assigned to other services. Many of these referrals will be hired with or without the OJT subsidy (Simpson, 1984a) and the others applied to JTPA only because their prospective employers demanded it.

EXHIBIT 3.1. DEFINING MAJOR TYPES OF BIAS IN DIFFERENTIAL IMPACT RESEARCH

Selection Bias

When participants who select or are selected into different program variants differ in ways that affect program outcomes, observed outcome differences between program variants could be produced either by program qualities or by participant characteristics. If participant characteristics are not taken into account, estimates of program impact will be biased. Since many such differences may exist but not be measurable, some degree of selection bias is always present in differential impact analysis.

Non-Response Bias

No survey locates and interviews 100% of those in the sample. If those who do not respond would have answered differently on average from those who do respond, then results of the survey will be biased in the direction opposite that of the non-respondents. This is also a case of methodological selection bias.

Non-representative (Censored) Sample

If some segment of the participant or employer population being studied were omitted from a sample, for example, participants who terminated without employment, conclusions generated from the sample would be biased in the direction opposite that characterizing the omitted segment. This is a case of selection bias by the analyst.

Confounded Program Variants

When two program variants are correlated with each other, the unique effects of each can be estimated if both are included in the same equation. However, if one is omitted, then the estimate for the included variant will absorb the effect of the omitted one, biasing conclusions in that direction.

Response Bias

If responses to interview questions differ systematically from true answers, those responses are biased. For example, if participants feel that the JTPA agency wants them to report success, findings from surveys identified as conducted by JTPA may be upwardly biased.

HOW CAN BIAS FROM NON-RANDOM SELECTION BE MINIMIZED?

Reducing the effects of selection bias follows the general logic of causal analysis. Each source of selection bias is an alternative explanation which can be countered only by inclusion in multivariate equations of variables which identify the selection process. The discussion below identifies four major sources of selection bias. For each, the aspects of differential impact analysis most likely to be affected and the measurement strategies most able to minimize the bias are indicated.

Sources of Selection Bias

1. **Legally eligible individuals may or may not apply to JTPA**, because of differences in information available, personality or motivational differences, or geographical differences in services available. This selection process is critical for net impact studies, but seldom biases differential impact analysis, which involve only comparisons among individuals already enrolled in JTPA. However, if this type of bias differs across SDAs, then statewide differential impact comparisons among SDAs will be affected.

Little protection from this type of bias is available to analysis which does not include an untreated comparison group. However, SDA level measures of program availability and participant measures of motivation for applying to JTPA may help assess possible differences between SDAs. In addition, standard demographic background characteristics are often correlated with motivational characteristics, allowing their inclusion in differential impact analysis to act as a partial proxy for direct measures of motivation.

2. After eligibility is determined, **the participant may or may not be enrolled into JTPA.** If the reasons are correlated with program outcomes, bias will result. The source of this type of selection may be:

- Program policies and practices such as targeting,
- Individual choice made after learning of program options, or
- Failure to locate a program placement of the type decided on for that participant.

Where differential impact analysis compares different JTPA service providers, this form of selection will bias comparisons in the likely event that different providers generate different selection.

Once again, measuring the source of the selection is the appropriate tool for combatting the bias. It is possible to measure policies intended to determine which eligible individuals are to be enrolled. In addition, measures of participants' demographic and work history characteristics may act as a proxy for agency selection or may indicate which participants best fit the agency's desired targets. Beyond that, the key measures of agency selection involve the proportion of eligibles for each provider who fail to enroll and the reasons why. Given some JTPA managers' reported emphasis on enrolling the most qualified participants, statewide differential impact analysis will be well advised to include agency level measures of intended and, where possible, actual selectivity by managers.

3. Aside from the decision to enroll in JTPA, **a participant may request a particular treatment**. If the reasons for that request also predict that participant's likely program outcome, the self-selection can bias estimates of how program activity affects outcomes. This is especially likely to occur where employers select desirable job applicants and then send them to JTPA to request OJT enrollment. Participant requests are most likely to involve a basic program activity, or a particular school or employer. Analysis comparing these most basic treatment variants is therefore the most likely to suffer from this source of bias.

The best protection against bias from self-selected treatment is measuring participants' route into JTPA: whether they requested particular services, and if so, which ones and why. One could also measure the degree to which particular service providers control the assignment to treatment versus allowing participants to elect their own treatment.

4. Participants may also be **assigned to particular treatments by program managers**. If treatment A rather than treatment B is assigned on the basis of factors which also influence program outcome, selection bias is present.

This source of bias has potentially pervasive effects on differential impact analysis. The reason is that rational service provider policy offers the most intensive services to those with the greatest need. That is, many JTPA services are intentionally compensatory. Since "greatest need" often translates to "least employable," the selection of services on the basis of need can bias estimates of treatment impacts on employment outcomes. For example, at termination the only participants who need job placement services are those who were unable to secure a job without assistance. Therefore, post-program employment success will appear to be higher among those who did *not* receive job search assistance.

This source of bias is difficult to mitigate. To identify compensatory effects of treatment, one must have measures of both the need and the treatment. Since these two factors have opposite effects, they cancel each other out and neither effect is visible without joint analysis of both variables. Even classical experimental design would have difficulty, since a necessary element of the treatment is the non-random assignment of services. Biased estimates must be prevented through measuring the selection process.

Measurement Approaches to Combat Selection Bias

When differential selection cannot be prevented, it must be identified by measuring the selection. Differential impact analysis has the advantage that it can attack from two different angles, using agency level measures and individual level measures. (See Appendix B for a discussion of measurement levels.)

Individual Level Measures have several advantages for combatting bias:

1. Ideally, they can include the agency's diagnosis of each participant's need and of its service prescription for each participant, as well as the treatment each individual received. Both the participant's true level of need and the agency's perception of each participant's need are important potential sources of selection bias.

2. They offer information on the explicit selection process by the JTPA agency. For example, we can learn how much intake time the agency spent with each particular participant. It is one thing to know what "full" intake includes (an agency level measure) and another to know that individual A received only a "fast track" intake review, while individual B was judged to require extensive pre-employment services.

3. They can include the route each specific individual takes into training. This proves to be one primary indicator of selection bias.

4. They allow precise measurement of program variants, increasing the power of the measures most important to any differential impact analysis.

Agency Level Variables also exhibit two particular strengths in combatting selection bias.

1. The problem of compensatory treatment cannot be fully solved by individual level measures, because no precise measures of need for assistance exist. Agency level measures indicate

resources *available* or provided *on average*. They are therefore much less influenced by compensatory treatment. For example, if Agency A provides job search assistance to only 5% of clients while Agency B does so for 40%, it is almost certain that many individuals of equal need will receive this service in Agency B but not in Agency A.

2. Agency policies directly affect selection. Targeting decisions, policy toward "creaming," policies regarding single versus multiple activity treatments, and the like, have some consistent effect on selection across all participants enrolled through a particular agency. Such agency policies can indicate selection on difficult-to-measure criteria such as how participants present themselves interpersonally.

TWO ANALYSIS GOALS CONFRONTING POTENTIALLY SEVERE BIAS

The most basic and consequential of program divisions, by service provider and by program activity, are the most likely to be effected by selection bias. Participants are more likely to know about them and to exercise choice regarding them. They involve basic resource allocation decisions and are therefore likely to be affected by geo-political concerns. Also, agencies are more likely to control which participants receive such basic services.

Different Service Providers

SDAs, and to a smaller extent their subcontractors, are located in different labor markets and political atmospheres. Although some of these differences can be accounted for through measures of the labor market environment and of agency policies, many will remain unmeasured. Therefore, some unknown degree of bias will persist in analysis across service providers, especially SDAs.

Basic Program Activities

Basic treatment options are designed in part to accommodate differences in participant needs and qualifications. In particular, job search assistance assumes job readiness, OJT assumes minimum acceptability to employers, and WEX assumes an absence of even the most basic job experience. Selection bias is likely to be especially serious in such cases, where it is explicitly called for. In addition, different treatments produce the outcomes mandated by JTPA through different mechanisms, making them complex to compare directly. (See Chapter 7.)

These concerns cumulate to the recommendation that where resources allow, sample size be large enough to accommodate separate analysis within each basic activity, along with tests performed across all activities. In addition, differential impact analysis should include either/or (dummy) variables indicating membership in the most common program activities. These variables will absorb basic program activity effects and also some unmeasured selection effects, thereby reducing bias in estimates of other effects.

Alternatively, only selected activities can be analyzed. This approach also has the advantage that some program variants of interest apply only within particular activities. For example, measures of employer characteristics apply only to employer-based treatments. This makes their analysis possible only when sample size allows activity-specific analysis.

PART. III

DEMANDS ON THE NUMBER OF TREATMENT CONTEXTS AND ON THE SAMPLE SIZE OF EACH

FOUR TYPES OF MEASURES

The intersection of two distinctions forms the four types of measures illustrated in Exhibit 3.2. One contrasts measures of individual treatment with agency level implementation variants. The second distinguishes between "either/or" measures of membership in specific treatment contexts and variables describing characteristics of programs or treatments. Each of these four types has somewhat different implications for sample structure, especially the number of treatment contexts to be compared.

EXHIBIT 3.2. FOUR APPROACHES TO MEASURING PROGRAM VARIANTS, WITH EXAMPLES

	Program Implementation, Once-Only Description of Service Provider	Measure of Individual Treatment
Either/or Measures Of Membership in a Specific Context:	Enrolled Through N.W. Corner SDA, Subcontractor No. 3 Versus All Others	Trained In Community College Program Versus All Others
Variable Descriptions Of Program Characteristics:	Percent of Services Performed In-House	Planned Length of Participant's Training Program

Specific treatment contexts may include enrollment through a particular SDA or service provider, training in a particular school, or participation in a particular program activity. In each case, the measure indicates simply whether the participant was enrolled in that particular organization or activity. This form of measurement is an effective way to *locate* impacts on post-program outcomes, but not to *explain* why they are located where they are. The case in which number of contexts is most at issue is when SDA or service providers are being analyzed.

Variable descriptions of program characteristics measure specific qualities which vary across all service providers or program activities rather than separating each as a whole from the others. This approach does not pinpoint concrete contexts within which differences occur. However, it helps explain *why* they occur, a quality which makes them especially helpful for program development. Knowing which qualities of SDAs to emulate may be more important than knowing which SDA performs best. Of course, utilizing both types of measures to gain both types of information is preferable.

The distinctions made in Exhibit 3.2 involve trade-offs in research design decisions. Measures of individual treatment require relatively expensive data collection tied to each individual participant, while program implementation variables measure only the smaller number of service providers. However, the ability to analyze implementation variants depends heavily on the number of treatment contexts included in the analysis.

Either/or membership variables are easily measured. However, their analysis demands an adequate sample size for *each* membership group. Analysis of variable program descriptions makes smaller demands on sample size, but require original data collection and demand that a larger number of different program contexts be included in a given analysis.

DEMANDS ON SAMPLE STRUCTURE MADE BY FOUR MEASUREMENT APPROACHES

Either/Or Membership Measures of Program Implementation or of Individual Treatment

Either/or membership measures place little restriction on the number of service provider or other treatment contexts required. If two contexts are identifiable, they may be compared by entering the dichotomous variable into an equation that also includes the appropriate control variables. We need only assess whether the adjusted average outcome is greater in one context than in others. This can be done by comparing two contexts or 20 contexts. Each comparison is a simple yes/no, so that 20 comparisons involve 19 yes/no variables (called "dummy" variables), just as 2 contexts involve 1 such variable.

This simplicity is gained at some cost. First, *explanatory power is no greater with 20 contexts than with two*. Each context is compared individually with all others. Second, the *reliability of such comparisons depends on the number of participants enrolled in each membership group.* Thus, no matter how large the total sample, estimates for membership in a category containing very few participants cannot be reliable. This means that analysis of this type may require large, disproportionately stratified samples. Membership in highly specific contexts, such as particular schools or employers, are usually immune to analysis because so few individuals belong to each context. (See Chapter 4 for a discussion of sample size.)

Variable Descriptions of Program Implementation

Variable descriptions of implementation involve a somewhat different trade-off. Since participants in each context receive a specific value on some measurement scale, the number in each context matters little. However, that advantage is purchased at the cost of requiring that multiple treatment contexts be included in the sample.

Imagine that two service providers have been measured on two variables: intensity of intake procedures and the degree of job search assistance provided. If we find that the two providers differ in outcome level, how can we decide which of these variables accounts for the difference? For that matter, how can we claim that either of these variables explains the difference? To assess whether intake or placement had the impact, we need to compare situations characterized by thorough intake but little job search assistance, and vice versa. But with only two organizations, that is not possible. These two agency characteristics, as well as any others one can imagine, are by definition perfectly correlated and cannot be disentangled. (In statistical terms, only one degree of freedom is available.)

This same problem faces research which compares more than two contexts but where the variable in question happens to differentiate only one from all others. An analysis reported by Franklin and Ripley (1984) illustrates. They report that program performance was lower in CETA prime sponsors characterized by "crisis management" style. While this finding appears reasonable, only one prime sponsor was so characterized, making their conclusion based on a comparison between 1 prime sponsor and 14 others. This means that any number of other qualities of that one prime sponsor could have produced the differences they observed.

In the case where three service providers are included in a sample, it is very likely that the problems discussed above will remain. However, there is now a possibility that, in unusual circumstances, one variable characteristic of service providers would have such a strong and consistent impact that a statistically reliable effect would emerge. The principle of parsimony -- using the simplest explanation consistent with the facts --becomes the guide to interpretation here. If the differences among outcomes in the three contexts fit well a single linear treatment measure, then it is parsimonious to explain findings with that one factor. If, however, they vary far from a linear fit, the less tidy but more accurate interpretation must be used, namely, that each unit differs from each other, for reasons we cannot demonstrate.

If we introduced a second treatment characteristic variable into the analysis based on three contexts, we would automatically revert to the case in which it is impossible to distinguish among competing explanations. In statistical terms, the number of variables which may be uniquely estimated may not be greater than the degrees of freedom, which equal the number of cases minus 1. Since these variables are measured only at the organizational level, the cases we are speaking about are the number of service providing organizations in the analysis.

Extending this line of thought, it is apparent that analysis including many SDAs or analysis of large SDAs including many service providers can be especially valuable for local program development. The larger the number of different agencies in the analysis, the more feasible are tests of agency implementation variables. More variables can be handled simultaneously and each is tested more reliably and less ambiguously. That is, other things equal, the more separate service providers are included in a sample, the lower the covariance among implementation variants is likely to be, strengthening the ability of multivariate analysis to estimate the unique effects of each.

This general rule, that the larger the number of contexts, the firmer the analysis of variable program characteristics, leads to a practical question: What is the minimum number of service providers required for a reasonable differential impact analysis of agency level implementation measures?

The answer is two-fold. First the bad news: the answer depends on many factors: variance in each independent variable, variance in the outcome variable, covariance among independent variables, and covariance between independent variables and the outcome variable. Therefore no precise minimum can be set forth. One might reasonably say that there is little point in pursuing analysis of variable program implementation measures with fewer than six or seven service providers. In many cases this would be too few, while in some cases, it would be sufficient.

Second, the good news: there is an analysis procedure which can in most cases protect against incorrectly attributing too much importance to variable descriptions of program characteristics. This procedure involves jointly testing *both* the variable program characteristic measures and the categorical either/or variables indicating enrollment in each particular service provider. Taking variable program characteristics which appear statistically reliable, we then add to the equation a set of dummy variables representing membership in each service provider. [2] If the variable program characteristics retain their statistically reliable effects, then our confidence in the initial findings remains high. If their effect in the equation is eliminated by the addition of the either/or membership variables, then we must conclude either that some service providers differ from each other but we do not know why, or that the initial test procedure was inappropriate. With small numbers of units, the latter is always a strong likelihood.

Variable Descriptions of Individual Treatment Variants

Individual level measures increase data collection costs but make fewer demands on the sample structure. The sample need include no minimum number of participants from any one treatment context unless the analysis wishes specifically to separate effects of individual treatment from those of implementation variants. Except where individual treatment happens not to vary *within* service providers, more individual treatment variables than agency level measures may be included in one equation and these variables do not place demands on the number of service providers in the sample. That is, individual treatment variants usually suffer less problems of colinearity than do the other types of measures discussed above.

Summary

Either/or membership variables may be tested comparing as few as two units, given sufficient sample size within each unit. In the case of variable descriptions of program implementation, state level (or other multi-agency) analysis is recommended. Variable descriptions of implementation can identify specific policy directions using economical agency level measurement. However, they become tenuous with small numbers of service providers. Individual level measures offer greater specificity, while suffering less from limits of number of service units or the number of participants in each.

---

[2] If all are entered simultaneously, only n-1 dummy variables may be included. If the analysis involves a forward stepwise procedure, n dummy variables may be included.

SECTION II.

RESEARCH DESIGN

CHAPTER 4

RESEARCH DESIGN FOR A PARTICIPANT FOLLOW-UP

Chapter 4. Research Design For a Participant Follow-up

Chapter 4 represents a transition from the conceptual issues discussed in Chapter 3 to more specific decisions about how to organize gross impact research efforts. The term *research design* is intended to include defining the population to be studied, developing a sampling strategy, defining a follow-up period, and selecting data collection methods for various types of measures. The recommendations here attempt to implement a high quality research design within the limits of data collection time and costs, and the ability to obtain valid measures.

PART I.

ALTERNATIVE LEVELS OF INVESTMENT IN THE ANALYSIS OF PARTICIPANT OUTCOMES

The first decision a state and its SDAs must make regarding program analysis based on follow-up data is whether the benefits are worth the costs involved. The elusive ideal would be to receive information useful to program management and development without adding data elements to an agency MIS or developing follow-up systems, and without having to locate additional financing to support analysis. Although that ideal is unreachable, it is possible to design cost-effective research and analysis at several different levels of investment.

This chapter will recommend design elements for three different levels of investment. These differ in the *research questions* they are able to address, the *specificity and power of measures* they include, the *protection from bias* that they provide, and their *cost*. It is important to offer these alternatives because of the relatively open-ended nature of gross impact analysis, because different states and SDAs will have different analysis goals in mind when making design decisions, and because funding limitations may eliminate designs which are preferable on other grounds.

THREE LEVELS OF INVESTMENT IN GROSS IMPACT ANALYSIS

The primary costs encountered by gross impact analysis are *initial setup costs, data collection costs,* and *data analysis costs*. The different levels of investment suggested below vary along each of these dimensions.

THREE LEVELS OF INVESTMENT IN ANALYSIS OF GROSS PARTICIPANT OUTCOMES

A. Differential impact analysis involving a minimal number of outcomes (post- program performance standards measures), agency MIS data, and agency level implementation variants.

B. Descriptive analysis of a broad range of participant outcomes.

C. Differential impact analysis involving a broad range of participant outcomes, MIS data, agency level implementation variants, and individual treatment variants.

Level A is minimal, involving no data collection on individuals, beyond that required under JTPA regulations. Post-program performance standards define the minimal participant follow-up measures. Descriptive analysis follows by definition, via annual status reports. Indeed, this design is recommended because it builds efficiently on the existence of post-program performance standards. In addition, differential impact analysis is also advisable.

One benefit of differential impact analysis of these data is that states could adjust SDA performance levels by reducing or eliminating the effects of individual background characteristics, to the extent that they are available in MIS files. Such adjustments would increase the comparability of SDAs within the state. Also, with the addition of a survey of service providers across the state, implementation variants can be analyzed, along with the basic program activities identifiable through MIS files. These tools can take states a considerable distance toward identifying advisable directions for program development. This in turn transforms the analysis effort into a solid basis for the state technical assistance role. In addition, agency level measures of selection can be included in the analysis.

Level B avoids costs of extensive analysis and of measuring treatment or agency implementation by centering on description of a wide range of participant outcomes. In so doing, it includes the most expensive form of gross outcomes data collection: interviewing recent participants. However, a large portion of the cost of such interviews is already required to measure post-program performance standards, making the addition of further descriptive measures less costly.

Level C Includes all participant-based data and analysis capabilities discussed in this guide. It is more costly than options A and B because it includes the collection of additional data elements on each individual in the sample, both during treatment and as part of follow-up interviews. However, its use of individual level measures of program variants gives it greater analysis power and specificity, and makes SDA level as well as state level differential impact analysis possible.

The remainder of this chapter is organized by design issues. Most recommendations are independent of which level of investment is intended. However, some design elements, especially those involving sampling, necessarily differ depending on which of the above options is chosen. In such cases, separate recommendations are given for each, along with a brief discussion of the rationale for each.

PART II.

IDENTIFYING THE POPULATION TO BE ANALYZED

The first step in designing a gross impact analysis is deciding which set of participants to include -- that is, how to define the *population* under study, the population to which conclusions will be generalized. Several issues must be discussed.

PROGRAM TITLES AND ACTIVITIES

The choice of JTPA Title depends primarily on managers' goals for the analysis. In addition, attention must be paid to the compatibility of measures across Title. Where the decision is made to include more than one title, some measures may need to be tailored to each, and separate sampling frames may be required.

If the goal is purely descriptive analysis, it is most reasonable to include the full range of program activities within whatever titles are selected. If the goal is an in-depth differential impact analysis, the selection of particular program activities may be preferable if managers wish to focus improvement efforts on them. Such decisions are entirely a matter of policy, not of research method.

PROGRAM STATUSES

The second decision regarding what population to study is whether the population should include all those found to be eligible for JTPA services, all those who were enrolled in JTPA, all statuses at termination, or only participants who were employed at termination.

All Eligibles versus All Participants

Some states or SDAs may wish to concentrate their analysis on program selection. This requires a population including all *eligibles* rather than only enrolled participants. This approach may be especially valuable if the gross impact approach is to be combined with a process analysis.

Studying all eligibles is not a design recommended here, because it addresses questions somewhat different from those this guide emphasizes and because of the increased data collection and storage costs. Such a study would, however, offer valuable tools for analyzing the selection process. It is essentially a separate analysis, the results of which could be an extremely valuable addition to a differential impact analysis. (See Grembowski, 1986.)

All Enrolled Participants versus All Terminees

This distinction is mentioned here to highlight the importance of including program dropouts, those terminating without completing the planned course of study or period of subsidized employment, in the population to be analyzed. In theory, all those who are enrolled must also be terminated, even if on the same day. However, if there is any possibility that these categories differ, it should be made clear that the study population is to include all those who were enrolled in JTPA for *any length of time*. In practice, this need not mean that participants are identified for the sample at enrollment rather than at termination. The distiction comes into play only when there is reason to believe that some individuals might appear among agency enrollees but not among terminees.

Termination Status

There has been a tendency for follow-up surveys to include only individuals who terminated with employment. In these cases, the outcome which can be measured is limited to "retention of the termination position." This approach reduces costs by allowing data to be collected through employer follow-ups and by selecting terminees most likely to remain residentially stable for the follow-up period. However, there are several reasons why *gross impact analysis designs should include all termination statuses* in the population to be studied.

- Post-program performance standards require follow-up of all termination statuses. If analysis is to be built onto that follow-up, inclusion of all statuses will occur automatically.
- Any estimate of a standard outcome measure demands follow-up of all participants. Aggregate measures are expressed in terms of averages or distributions -- average wage, proportion at very low wage rates, proportion employed, and the like. Any of these measures would be badly inflated by the exclusion of the group least likely to be employed -- those unemployed at termination. On the other hand, to estimate this group's employment and earnings at zero (their status at termination) would seriously underestimate program success. In short, accurate estimates of program outcomes at any given point in time require measuring all those enrolled.
- Including those not employed at termination is the only way to capture delayed employment, which can be especially problematic for classroom training in fields where openings fluctuate seasonally or with large contracts.
- It also helps establish which participants have left the labor market, which are experiencing prolonged unemployment, and the reasons explaining delayed employment.
- It makes descriptive findings more comparable to those based on Unemployment Insurance data, another useful source of data on JTPA outcomes (Johnson, 1986).
- It protects against differences in service providers' methods of defining terminations.
- One value of follow-up research is to help assess why some individuals were not assisted by program participation. Omitting these people from the population studied eliminates that value.
- Service provider cost estimates are based on averages or on combined totals for all participants. To be compatible, outcome estimates must also be based on all, or a sample of all, participants. If costs for all were compared with outcomes of only the most successful, estimates of cost effectiveness would be seriously inflated.

There is one less inclusive follow-up design, discussed in Chapter 5, which can be mentioned as a less desirable, but much less expensive option: an employer-only follow-up of those individuals employed at termination. This design is less desirable than a follow-up of all participant statuses because it omits *two* follow-up issues: late employment and turnover after having held a termination job. However, if either of these is to be omitted, it is preferable to omit both, thereby focusing on an analytically clean question: what is the sheer job retention rate produced by the JTPA program? This means that analyses involving only employer follow-ups can adequately measure this one participant outcome as a side-benefit of the employer analysis.

DURATION OF THE STUDY

While it would be convenient to concentrate data collection into a few months, this shortcut would endanger the validity of the research, introducing some known biases and others less easy to identify. The population to be studied should therefore be defined to include *all enrollees or terminees throughout the full year.* There are three major reasons why any period less than one year could bias results:

First, seasonal labor market variations will affect outcomes, depending on the industry and occupation in which participants are trained and placed.

Second, in the case of classroom training, some institutions tend to end courses during particular months. A time frame which included the scheduled completion month but excluded others would produce interviews with a disproportionate number of program completers. With different months, the opposite bias might occur. Since completers are more likely to gain employment, estimates of all outcomes from CT can be easily biased in this way.

Third, different service providers develop different policies concerning when to commit their funds. In particular, some engage in "front-loading" to insure that they will meet their planned enrollment levels, and some develop strategies specific to local seasonal fluctuations. These policies influence the mix of program placements and selectivity available to service providers. For example, CT is easier to enlarge or decrease at some times of the year, and employer-based interventions during others. Year-round data collection avoids biases from these practices, and also enables JTPA systems to test whether such practices affect program outcomes.

Because of these rather serious bias issues, the recommendation of this guide is that *data be collected year round rather than during a concentrated period.* Once again, those cases in which analysis is to be integrated with the measurement of post-program performance standards will necessarily cover the entire program year.

It is possible to partially recoup the loss of efficiency from the year-round design by concentrating interviews in short bursts. The implementation requirements for performance standards limit flexibility in this regard. Greater leeway is possible for follow-ups which are separate from performance measures.

The second apparent cost of structuring data collection year-round is the delay between the beginning of data collection and the completion of the analysis. If data cleaning, data set merging, analysis, and writing take about six months, a six month data collection scheme produces a report in one year while a full year data collection does so in 18 months. There is no shortcut here, although analysis programs can be developed and debugged ahead of time, and partial findings can be generated for the early data collection period. However, for states or SDAs that continue program analysis on an on-going basis, this problem diminishes after the first year. If data are entered in a compatible, continuous-update fashion, then analysis can be performed covering any period involving multiples of 12 months.

SUMMARY OF RECOMMENDATIONS CONCERNING THE POPULATION TO BE STUDIED

1. Program Titles and activities must be selected on the basis of policy goals.
2. Within Titles and activities chosen, all participants enrolled for any time should be included.
3. Similarly, participants of all termination statuses should be included.
4. All enrollments or terminations during a full 12 month period should be included.

PART III.

DESIGNING THE SAMPLE

SAMPLE SIZE AND ESTIMATED SAMPLING ERROR

Once a population is selected for analysis, the question becomes how large a sample should be in order for calculations to estimate accurately the patterns within the entire population of participants. Sample size is critical to the usefulness of any analysis project because it dramatically affects the margin of error -- the reliability -- of conclusions. A conclusion that one program variant has 10% higher retention rate than another means little if the margin of error for that estimate is 20%. Assuming a representative sample, the primary determinant of error margin is the number of cases upon which estimates are based. Therefore, the first decision must be how many cases are needed in order to generate a level of error acceptable to those who will use the analysis results.

One essential reason that survey research has become such a widely used method is that the accuracy of estimates rises rapidly as we move from very small samples to samples of modest size, yet samples of modest size are nearly as accurate as very large samples. The reason is that error decreases as a function of the square root of the sample size. More precisely, the estimated error associated with any measure depends on the *standard error* of that measure. (For random samples only, the standard error equals the standard deviation of the measure divided by the square root of the sample size -- i.e., se = sd ÷ sqrtN). Thus, for example, if an income measure has a standard deviation of $4,000, the standard error is about $800 with a sample of 25, $400 with a sample of 100, $200 with a sample of 400, and $100 with a sample of 1600.

One can see the danger of relying on a very small sample. However, it is equally evident that the marginal improvement from each increase in sample size is quickly reduced as sample size becomes larger. Adding 375 to a sample of 25 reduces error in our example by $600. However, another 1200 cases would be required to trim a further $100 off the standard error.

When samples are large (over 100), the standard error tells us with precision how wide our margin of error is, within a given level of probability. This margin is referred to as a confidence interval. If we wish to be 95% confident that the true value of some outcome is within our confidence interval, that interval must include the values from two standard errors below our sample estimate to two standard errors above the estimate. Nineteen of 20 times, 95% of the time, the true value will lie within that margin. It will be within one standard error 68% of the time, and within three standard errors 99.9% of the time. Thus, in the example above, if our sample is 400 and our standard error therefore $200, our 95% sampling error is twice the standard error, or $400. If our estimated average income is $6,000, we may be 95% certain that the true average is between $5,600 and $6,400.

Working backwards, we may use these same calculations to determine what sample size we *want*. As stated above, se = sd÷sqrtN. Therefore, $N = (sd \div se)^2$. That is, the desired sample size equals the square of the standard error of the measure in question divided by the standard error we are willing to accept. For example, let us say we want a maximum earnings sampling error of $200 with 95% confidence, making the desired standard error $100. If the earnings measure has a standard deviation of $4000, we may calculate $N = (4{,}000 \div 100)^2 = 1600$. Therefore, given the error level we desire, a sample of 1600 completed interviews is required. The first step in determining proper sample size must be establishing the desired margin of error. After that point, a target sample size is derived mathematically.

The difficulty is that *we do not know before the research begins what the standard deviation of our measures will be*. Therefore, we cannot calculate a precise standard error or sampling error. Two approaches may be taken to solve this problem. First, one may consult previous research which used the same measures and use the standard deviations generated by that research to determine sample size. Second, recognizing that different types of outcome variables will generate different standard deviations, one can base sample calculations on the type of measure which is in general least reliable -- the dichotomy[1]. If the sample size is satisfactory for that type of variable, it will be satisfactory for all. For the reader's convenience, Appendix B supplies examples of the sample sizes required to produce given sampling errors with 95% confidence, for a dichotomous variable. It is shown there, for example, that if we estimate that job retention (a dichotomous yes/no variable) is likely to be about 70%, and we want a sampling error of no more than 4%, a sample of just over 500 is required.

Planners will be well advised to consult one of several thorough texts on sampling (e.g., Kish, 1965; Sudman, 1976) or to employ a sampling specialist in cases where sampling appears problematic or in order to determine the most cost-efficient sample.

LEVEL OF INVESTMENT AND SAMPLE SIZE

Each of the three levels of investment in gross impact analysis identified at the outset of the chapter makes somewhat different demands on its sample. Possible adjustments to sample size are therefore discussed for each level.

Level A. Minimal Differential Impact Analysis With Implementation Variants

This approach places special demands on the service providers included in the sample. Where either/or membership variables are being analyzed, the question is whether *each membership group* is large enough to produce reliable findings. (See Chapter 3.) For example, if the analysis asks whether any of several SDAs differs significantly from the others, each SDA must be represented by a subsample large enough to allow a reliable estimate of the outcome within that SDA. Sample stratification is often recommended in such cases, as discussed below. When the choice is to stratify, sample size usually increases since the purpose of the stratification is to achieve adequate samples of small units. Where variable descriptions of program variants are analyzed, the question is whether the sample includes a sufficient number of different service providers, rather than a given number of participants from each.

In either event, the question when planning sample size is more how the sample is to be internally composed, than how large the overall sample should be. Where samples are to be stratified, they should be constructed from the SDA level up, rather than from the state down. Proposed requirements for implementing post-program performance standards are perfectly suited to this approach, since they are statewide yet specify sample sizes for each SDA.

Level B. Descriptive Analysis of a Range of Outcomes

If the only goal of an SDA's follow-up study is the accurate estimation of one year's program outcomes, then for small SDAs, the *size of the population* (the total number of participants enrolled), as well as the *size of the sample*, influence the margin of error. The adjustment is typically small, but becomes noticeable where the sample represents a large proportion of the population. To find the *corrected sample size* (CSS) which will produce a given error margin, you must first find the uncorrected sample size (USS) for large populations, as outlined above. Then calculate what proportion of the population (POP) that sample represents, and square that value (calculate $(USS/POP)^2$), as shown in the formula below:

$$CSS = USS * (1 - (USS \div POP)^2)$$

---

[1] A dichotomy is a variable taking only two values, such as gender or whether one is employed. Some continuous variables also have large standard deviations. The more homogeneous the population is thought to be, the smaller the sample can be. If we wish to study heterogeneous populations such as JTPA, we must pay the price of a larger sample.

For example, assume that we have a total of 600 participants in the population to be analyzed, and that we have decided that we want to keep our error estimate around the 5-6% level, for a 95% confidence interval. Appendix B dictates a unadjusted sample size of about 300. Thus:

$$CSS = 300 * (1 - (200 \div 600)^2) = 300 * (1 - .11) = 267.$$

Concrete examples of these adjustments appear in the technical assistance supporting the JASR post-program data collection requirements.

Level C. Differential Impact Analysis With Individual Treatment Variants

All differential impact analysis makes greater demands on the sample than when the analysis is entirely descriptive. Depending on the *number of variables being analyzed, their distributions* and *their intercorrelations,* reliable findings demand somewhat larger samples than are required for descriptive analysis.

There is unfortunately no hard and fast rule about how to adjust sample sizes to make such internal comparisons more reliable, except that the sample size must not be reduced. Analysts typically leave the sample unadjusted or increase it by 20-30%. The larger the sample, the more reliable the analysis. Therefore, costs often dictate how much the sample is enlarged. Sudman (1976) offers a rule of thumb which may be used: each category in major breakdowns (e.g., basic program activities) should contain *at least* 100 members, and each category of minor breakdowns (e.g., categories of employer size or training skill area) should contain at least 20 members. Such a rule of thumb is not intended as a requirement, but as a diagnostic tool. If many of the comparisons planned would fall short of these category sizes, the analyst should consider increasing sample size.

WHEN SHOULD SAMPLES BE STRATIFIED?

Populations are sometimes divided into subgroups, or strata, each of which is sampled separately. Strata may be sampled in proportion to their numbers in the population, or disproportionately. Although it is sometimes believed that samples must be proportionately stratified to insure the proper number of members with various background characteristics, this belief is in error. Proper sampling procedures insure a representative sample. No reason exists to consider *proportionate* stratification in gross impact analysis.

In general, disproportionate stratified sampling should also be avoided except where the main goal of the analysis is the comparison of service providing organizations, as in the Level A investment discussed above. Disproportionate stratification involves additional administrative oversight, and a process of reweighting the finished data set in order to produce accurate descriptions of the entire population of participants, thus increasing costs and probability of error. Such samples usually should not be analyzed without expert guidance.

There are several conditions under which disproportionate stratified samples are sometimes recommended, but only one of these is applicable to gross impact analysis (Sudman, 1976). However, that one reason is central to statewide differential impact analysis as well as to analyses of subcontractor performance, conducted by large SDAs. The need for disproportionate stratification of gross impact samples arises when the analyst's emphasis is on comparing or reliably characterizing subpopulations rather than on characterizing the entire population of participants. This occurs in the case of post-program performance standards, where welfare recipients are treated as a separate stratum. In addition, statewide analysis aimed toward comparisons among SDAs, or SDA level analysis comparing service providers should consider stratifying to insure reliable characterization of smaller units.

Level A. Differential Impact Analysis With Implementation Variants

When implementation variants are to be analyzed, there is little reason to consider sample stratification. Comparisons are made among the different states of each variable, not among particular organizational units. However, when the major focus of the research is the comparison of either/or membership categories such as enrollment through one particular service provider, the size of the total sample matters less than the size of the sample from *each* category. (See Chapter 3.) If I have a sample of 1000, but I am comparing one SDA including only 50 participants with all others, I am saddled with two error margins: about 3% for the sample of 950 and about 12-14% for the SDA with a sample of 50 participants. This means that the estimated difference between them must be at least 15% before we can reliably draw the conclusion that they differ. Comparing two equal groups of 500 requires about half that difference for reliability.

This constraint represents the primary cost of differential impact analysis aimed at SDA level comparisons. What may appear to be a study of "the state" is in some regards a combined study of many separate service providing organizations. Therefore, *when state level differential impact research efforts are being mounted with the primary goal of reliably estimating differences among SDAs, disproportionate stratification of the sample becomes advisable.*

Such samples should be planned from the SDA level upwards. First, planners should establish what error they can live with when estimating values for each SDA, and calculate the minimal SDA sample size on that basis. In this instance, requirements for measuring post-program performance standards have provided a first approximation which is likely to be satisfactory for most states. Each SDA is required to complete a number of interviews sufficient to produce a 95% confidence interval around SDA estimates of approximately 5%. Combining those samples statewide would construct a disproportionately stratified sample sufficient for firm comparisons statewide and among SDAs.

If smaller error is desired, or if service providers within one SDA are the units of interest, the same approach can be taken, but the sample would need to be enlarged or stratified by service provider. Except for very large SDAs, any single SDA analysis with the goal of generating reliable comparisons among service providers must probably consider a brief follow-up of all, or nearly all, participants.

Level B. Descriptive Analysis Only

Sample stratification should not be entertained for descriptive analysis unless there is a need to increase reliability when separately describing some specially designated subgroup. For example, if an SDA had a special interest in observing outcomes of a small pilot program, it might wish to include all participants from that program in the sample.

Level C. Differential Impact Analysis With Individual Treatment Variants

When measurement of treatment variants occurs at the individual level, stratifying the sample by service provider is no longer necessary. Doing so is likely to offer only minor improvements in statistical power, while at the same time distorting the variance among treatments within the state or SDA and forcing adjustments to both descriptive findings and to the results of differential impact analysis. In particular, findings based on unadjusted stratified samples will efficiently test differences across strata, but may not be generalizable to the full population from which the sample was drawn. To be generalizable, the sample must be *representative* of the population to which findings are to be generalized.

When states wish to conduct differential impact analysis of individual treatment variants, but also wish to integrate data collection with the stratified samples which are required for measuring post-program performance standards, analysts should report both descriptive findings and also results of complex models after reweighting the sample. (See Chapter 6 for an example of reweighting samples.)

States or SDAs who decide to identify samples at enrollment and to use those samples to satisfy the DOL post-program performance requirements should keep in mind one caution. The sample size required by DOL for each SDA's sample is based on terminations during four quarters. Although DOL allows sample identification at enrollment,[2] that approach creates a lag time problem because many participants will not terminate during the same quarters they enroll. Therefore, the sample identification and tagging process must precede measurement by a time period approximately equal to the longest enrollment period allowed. In addition, some participants will be transferred between programs, or their paperwork delayed, delaying their termination longer than envisioned and reducing the planned number of terminations during any one year. Therefore, samples identified at enrollment should be somewhat larger than the minimum sample required by DOL.

SUMMARY OF RECOMMENDATIONS CONCERNING SAMPLE SELECTION

1. For **descriptive analysis**, sample size should be established on the basis of the error margin policy makers find acceptable. Standard deviations used in calculations can be estimated from prior research or based on dichotomous variables. Small SDAs may correct their sample size downward, using the formula shown.

2. **Differential impact analysis** increases demands on sample size, making modest increases in sample size advisable.

3. In one case, that of **state level differential impact analysis comparing SDA outcomes**, it is recommended that disproportionate stratification be used to equalize SDA subsamples.

4. Decisions regarding sample size and disproportionate stratification procedures should be based on expert guidance.[3]

PART IV.

SAMPLE SELECTION

SAMPLING PROCEDURES

The sample of participants who are interviewed must be *representative* of the population being studied. *None of the claims for sample efficiency or reliability hold when samples are not representative.* Samples are representative when *each element of the population has an equal chance to be included* in the sample. Sample selection procedures must guarantee equal probability of inclusion, eliminating any purposeful or accidental selection. The classical way of doing that is to select each individual from an ordered list, using a table of random numbers. However, two more convenient methods are equally valid.

---

[2] This statement is based on draft documents rather than final DOL publications and should therefore be confirmed.

[3] Statewide comparisons among SDAs can adapt DOL requirements for SDA level sample size to guide disproportionate stratification by SDA.

First, the last three digits of participants' social security numbers are random with respect to any meaningful characteristic of individuals. Therefore, sample members may be selected by identifying a range of three digit numbers which would produce the required sample size, and including all participants with numbers falling in that range. If, for example, 25% of names are to be sampled, the lower end of the range is chosen at random and the upper is set at 250 higher. The second method is systematic sampling based on a random start. If 25% of the population is to be sampled, a list of names is prepared and one of the first four is chosen at random. Then, every fourth name is included in the sample.

These methods are easy to implement, with or without computer assistance, yet produce the representative samples required to prevent biased sample selection. Each method is well described in the technical assistance guides for post-program data collection, which accompany the revised JASR Reporting Requirements.[4]

To prevent any question about the care with which this critical step was implemented, analysts would be well advised to keep on file the original population, the completed sample, and the procedure used to identify them.

RESPONSE RATE AND SAMPLE SIZE

Once calculations have been performed to establish the number of cases desired for analysis, we move to the process of identifying specific individuals for inclusion in the sample. The first step is establishing the number necessary to generate the desired sample of *completed* interviews. This number is calculated by dividing the desired number of completed interviews by the planned survey completion rate. For example, to complete 400 interviews at a completion rate of 70% would require that 571 (400÷.7) names be identified in the initial sample.

WHEN SHOULD PARTICIPANTS BE IDENTIFIED FOR INCLUSION IN THE SAMPLE?

Participants may be selected at either *program entry* or *termination,* as long as the full population of participants is available for the sample. From the practical research administration viewpoint, it is preferable to identify the sample at termination for investment levels A and B. The follow-up period must be defined from the point of termination, and termination employers must be identified at that point, making it convenient to identify the sample then also.

There are three conditions under which it is preferable to identify participants at their program entry. First, a full differential impact analysis (investment level C) requires the measurement of individual treatment, begininning at enrollment. Rather than collecting this information on all participants, an agency sampling a relatively small proportion of total enrollments would find it most efficient to identify participants for inclusion in the analysis *upon entry*. At that point, they could be specially tagged for the collection of individual treatment data and for inclusion in the post-program follow-up.

Second, some states or SDAs may wish to perform an inclusive analysis of program selection, and to base their sample on all *eligibles* rather than on enrolled participants.

Third, if there is any slippage in record keeping, so that some enrollees might no longer be identifiable at termination, the sample should be identified at the point of enrollment.

---

[4] At this writing the technical assistance guide for post-program data collection is in draft and is not available for citation.

SUMMARY OF RECOMMENDATIONS CONCERNING SAMPLE SELECTION

1. The number of names identified should equal the desired number of completed interviews divided by the expected completion rate.

2. In most cases, sample members may be selected at termination, the most convenient point in time.

3. Sample members should be identified at enrollment if agency records of individual treatment will be kept during enrollment or any doubt exists that all enrollees can be identified at termination, and if data collection is separated from requirements for post-program performance standards.

4. Random sampling using a range of social security numbers or systematic sampling with a random start are recommended for choosing specific sample names.

PART V.

WHAT FOLLOW-UP PERIOD IS RECOMMENDED?

The number and timing of follow-up efforts is dictated primarily by resources and analysis goals. A major element in the resource equation is the requirement that three performance standards be measured by interviewing participants at 13 weeks after termination from training. A large portion of the costs of any survey occur before the first question is asked (recording locator information, identifying a sample, keeping records on that sample, tracking hard-to-locate former participants, hiring and training interviewers, setting up interview phone banks, multiple calls to locate the participant). Therefore, the most logical analysis design from the standpoint of data collection efficiency is to add questions to the required 13 week follow-up survey. In addition, the issue of selecting a follow-up period should be examined on its own merits.

Follow-up surveys are subject to serious *sample attrition* if the first or only interviews are conducted very long after termination. Since sample attrition introduces unknown biases, it is *preferable to conduct shorter term follow-ups and to achieve higher completion rates.* Those planning longer term follow-ups will be well advised to build in intermediate follow-up interviews as well. However, recent studies testing how well various follow-up measures predict long term net impact of JTPA find three month follow-ups much stronger than termination data alone, six months stronger than three months and nine months stronger than six months (Zornitsky, et. al., 1985b; Geraci, 1985). While the gain from each additional delay is smaller than the one before, each does offer improved reliability.

These considerations, taken with the requirement that a three month follow-up be performed, recommend two options for follow-up analysis:

1. That data collection for analysis purposes be combined with measurement of performance standards at three months, and

2. That a longer term follow-up be conducted using the sample of individuals interviewed at three months to satisfy performance standards.

THE BASIC THREE MONTH FOLLOW-UP

The option of conducting all follow-up at three months carries obvious financial advantages. In addition, it appears a reasonable approach, for several reasons.

- Three month follow-up data offers marked improvement over termination data alone in predicting long term program impact.

- A three month delay is long enough to allow the rate of employment after classroom training to stabilize. It is also long enough for OJT placement to stabilize after the post-contract drop-off, even where 30 day delayed performance payments may delay that drop-off. (See Appendix D.)

- Three months is long enough to eliminate most inconsistencies introduced by the tendencies of some service providers to make more extensive use than others of the "administrative hold for job search assistance" category following the program.

OPTIONS FOR EXTENDED FOLLOW-UP

Multiple Surveys

Systems wishing to conduct longer term follow-ups may employ essentially the same procedures as with a three month follow-up. If multiple contacts are envisioned, e.g., after three and nine months, the sample and data sets are intact after the first follow-up. Although two surveys cost more than one, the cost is not doubled. Little additional setup is required for the second survey; it can be quite brief, adding moderate phone and interviewer costs; and less tracking time is typically necessary for the second follow-up.

Two stage follow-ups must draw larger initial samples than would be drawn for a short term follow-up alone, in order to plan for attrition at each stage of data collection. For example, if the interview completion rate is 70% at a three month follow-up and 80% of those at a six month follow-up, the six month sample represents only 56% of the original sample. If sample sizes are established at the minimum allowed for post-program performance standards, they will be too small for useful analysis at six or nine months.

One option for longer term follow-ups deserves special consideration because it produces a full follow-up of participants and of termination employers at only a slight increase in cost over a participant-only design. It is a stepwise combination of employer interviews followed by participant interviews where necessary. (See Chapter 6.) Since DOL requirements for the 13 week follow-up stipulate participant contact, this design is recommended for longer term follow-ups.

Unemployment Insurance or Other Official Data

As an option, states or SDAs planning to do longer term follow-up may wish to consider using Unemployment Insurance wage records if they are available. Once the UI data base is arranged, a one or even two year follow-up is as easy to perform as a six month follow-up. Use of the UI system is detailed in Johnson's (1986) net impact model.

However, one factor limits the usefulness of UI data as a *gross* impact measure: UI data cover only individuals who maintain residence within the state. In the net impact approach, movement out of state is assumed to be equivalent for treated and untreated groups. However, gross outcomes are measured only for participants, making movement out of state a serious problem. One cannot determine whether a record of zero earning represents continuous unemployment or movement out of the state.

Use of this approach is recommended only if a separate tracking effort is mounted for those individuals with zero UI income, to estimate the proportion who moved out of the UI reporting area. That estimate could then be used to adjust estimated job retention rates.

DEFINING TERMINATION

One other issue must be resolved in order to define the period of follow-up: the point of program termination. What is the "program end point" from which individuals are to be followed? The most common approach is use of official program termination date. For most research, this is the only date available. However, for in-house program analysis, the completion of program activities defined in the training plan or participant contract may be a more meaningful termination date.

The second definition differs from the first in that large numbers of participants are placed in an informal job search assistance category for a period after their training. This extended period is ambiguous with respect to services provided. For some program activities, CT in particular, it is reasonable to define a post-program job search period as a necessary element of the intervention, a fact recognized by formal JTPA policy. In other cases, clients are placed in an "administrative hold" status primarily as a hedge against low termination statistics.

Despite these ambiguities, the date used to calculate the follow-up period should be *official termination.* This is the termination point defined for post-program performance standards and it recognizes the validity of offering job search assistance after training. Also, if the end of services were used as the beginning of the follow-up period, the administrative hold option along with paperwork delays would lead a considerable portion of the sample to be identified too late to satisfy a three month follow-up time criterion.

Nevertheless, a good analysis must measure an accurate time line for each participant, including dates when program activities ended and when employment began. The recommendation here is therefore three-fold:

- First, for purposes of calculating when to conduct follow-up interviews, termination should be defined as official termination of JTPA enrollment.
- Second, job history information, including employment dates, should be gathered for the entire period beginning with completion of training or contract.
- Third, dates should be recorded indicating the beginning and end of each official treatment and of an administrative hold period, where applicable. These dates and the pre-termination job history could be recorded efficiently during a termination exit interview.

SUMMARY OF RECOMMENDATIONS CONCERNING FOLLOW-UP PERIOD

1. A single follow-up or the first of multiple follow-ups should usually occur after three months, in conjunction with measurement of post- program performance standards.
2. Extended follow-ups can be added using the same three month follow- up sample, provided that the initial sample size was enlarged to plan for the additional attrition.
3. A combined employer-participant follow-up is highly efficient and is recommended for follow-ups not being coordinated with the measurement of 13 week post-program performance requirements.
4. The follow-up period should be calculated from the point of official JTPA termination. However, employment outcome measures should extend from the close of program services.

PART VI.

DATA COLLECTION METHODS

Gross impact analysis involves data collection through follow-up surveys of participants and employers. Chapter 1 briefly reviewed some reasons why surveys can be an especially valuable tool for analysis of JTPA outcomes. In addition, the rapid expansion of the survey research industry has been accompanied by a growing literature on its strengths and weaknesses (e.g., Rossi, Wright and Anderson, 1983). A brief review of selected literature on survey research is included in this guide. (See Appendix B.) A very brief summary of points covered there is included here in Chapter 4.

MEASURING OUTCOMES AT FOLLOW-UP

Reliability and Validity of Surveys

- We have learned enough about asking questions to make surveys quite reliable when properly constructed.

- *Response bias,* one threat to the validity of surveys, appears not to cause serious problems as long as "loaded" questions are avoided. There is, however, a tendency for respondents to inflate reports of employment or income slightly.

- Numerous *non-response biases* are almost certain to exist in any survey. However, these are unlikely to pose a serious threat to well-conducted surveys. Some known biases push estimated outcome levels upward while others depress estimates. Still others are unknown, but are normally small.

- Several tools are available for combatting non-response bias:

    1. The higher the completion rate, the less a given non-response bias affects estimates based on completed interviews. Accurate locator information collected by JTPA agencies and techniques for tracking participants who are difficult to locate help raise the completion rate.

    2. Interviews should be concentrated during times when a representative cross-section of participants are available: evenings are best, with weekends acceptable and a mix of callback times advisable.

    3. Equal effort should be applied to locate participants from different geographic areas, enrolled through different service providers, etc.

    4. The degree of non-response bias should be identified by comparing termination statuses and MIS characteristics of responders and non-responders. In rare cases, after-the-fact adjustments to the sample can reduce specific non-response biases.[5]

---

[5]. Sample adjustments for gross impact analysis differ somewhat from those required for post-program performance standards. In the latter case, adjustments are made to the final aggregate annual status report. For analysis purposes, the sample must be weighted beforehand, so as to simulate a sample without non-response bias. Adjustments should leave the sample size unchanged, to minimize effects on statistical tests.

However, sample adjustments are no substitute for careful tracking and high response rates. Adjustments duplicate the characteristics of individuals contacted. These may differ from those not contacted. Indeed, adjustments based on few individual characteristics may introduce more severe bias than they eliminate.

Telephone versus Other Surveys

- In-person interviews are far more expensive than telephone surveys and roughly equal in terms of validity and reliability.

- Mail questionnaires are generally less expensive than telephone surveys, but are subject to serious non-response bias. They also involve extensive hidden costs in paperwork and time delays. Where phone calls are local, costs of phone and mail surveys are more equal.

- For JTPA follow-up surveys, telephone interviews are clearly the preferred method.

MEASURING PROGRAM VARIANTS

For those who plan to engage in differential impact analysis, program variants must be measured as well as outcomes. The method by which they may be measured differs for agency *implementation* measures and *individual treatment* variants.

Implementation Variants

Variables characterizing how each SDA and other service providers implement JTPA services must be measured at the organizational level. That is, measurement of implementation variants involves a survey of service providers. The measurement instruments used to characterize each provider must be constructed by some central agency such as the state, or by consensus among agencies. Whether service providing agencies are surveyed in person, over telephone, or through mail is less important than that responses accurately characterize each agency. Specific measures are discussed in Section III of this guide.

Individual Treatment Variants

Variables characterizing the treatment services received by each participant must be measured specifically, with each individual as the unit of observation. The most reliable and cost-efficient method for collecting such data is to include selected, standardized measures as part of each participant's agency record. These measures would be recorded by program officers at the appropriate points in each participant's movement through the system. For example, intake services would be recorded at the completion of intake; program activity along with its scheduled length and cost would be recorded when the contract is finalized; multiple services would be recorded as they occur; and so forth.

Measures should be selected so as to be reliably knowable by program operators, and such that key information on the treatment process can be obtained without overburdening staff time. Much information useful for differential impact analysis usually exists in relatively inaccessible form in agency records or contracts. Unless systems have moved to full computerization of program records, it is simplest to record these treatment variants on specially constructed forms, throughout the treatment process. Other information is available only from participants or from the program officer. All these sources are readily available to program officers.

An alternative form of data collection is available for some, but not all, standardized treatment measures. If these are not collected by service providers, either by design or because service providers resist the request of SDA or state administrators, selected measures can be taken during follow-up interviews. This approach is a second best design compromise, because it is somewhat more costly, it is somewhat less reliable, and it excludes some variables. However, it does make differential impact analysis possible if the preferred method is not available.

MEASURING CONTROL VARIABLES

Some of the variables collected to guard against error from selection bias are program variants, collected as discussed above. Other control variables include the various individual background characteristics and pre-program work history measures, which may predict post-program outcomes irrespective of program intervention. Many of these are routinely included in MIS files. Others, such as precise pre-program versions of post-program outcome measures, demand original data collection. Where service providers collect data on individual treatment variants, these could be included as part of the data gathered at intake.

The third category of control variables measure aspects of the demographic and labor market environment, gathered from published sources.

CHAPTER 5

RESEARCH DESIGN FOR AN EMPLOYER FOLLOW-UP

Chapter 5. Research Design For An Employer Follow-Up

Chapter 5 should be read after Chapter 4. Many design elements are common to both employer and participant studies. These areas of commonality are discussed only in Chapter 4. For ease of cross reference between the two chapters, topics are covered in the same order, and whenever possible using the same headings, in the two chapters.

ALTERNATIVE APPROACHES TO THE ANALYSIS OF EMPLOYER OUTCOMES

Different goals can direct the analysis of JTPA employers. Each generates design variations, mentioned when appropriate throughout this chapter. To set the stage, basic variations are listed below. One set of variations is created by defining which employers are to be interviewed, the other by defining analysis goals.

Which Employers Are Analyzed

- Analysis of **placement employers**, all those who employed participants at their termination from any activity, but no others.
- Analysis of **participating employers**, all those who participated in the delivery of services, regardless of the termination status of the participants involved.
- Analysis of both placement and participating employers.

Analysis goals

- Descriptive analysis of employer outcomes only.
- Differential impact analysis using only employer outcomes or participant outcomes measured through employers.
- Differential impact analysis using *integrated* employer and participant data sets to analyze the full range of employer and participant outcomes.

Integrated Design Options

To some extent, differences among analysis goals overlap with different definitions of the employer population being studied. In particular:

- Minimal **descriptive analysis of employer outcomes** most reasonably calls for a representative sample of placement employers.

- Limited **differential impact analysis of employer outcomes** can be performed using either termination employers or participating employers. However, since employer costs and benefits are quite different for these two groups, such analyses should be performed separately.

- **Differential impact analysis of participant outcomes** can be greatly enhanced by integration with data from participating employers. This approach would involve measures of the services provided by employers as well as of benefits at outcome. Data from termination employers, on the other hand, can add only marginally to information available in participant interviews.

IDENTIFYING THE POPULATION TO BE ANALYZED

Program Titles and Activities

The most basic of all employer design questions is whether the population being studied includes all termination employers, all participating employers, or both. In addition, managers who wish an in-depth analysis of one specific program activity may prefer an even more specific definition, such as all OJT employers. Aside from modest differences in cost, these decisions should be made on the basis of policy objectives: to which programs do managers wish to apply the results? Are specific services earmarked for development? Is descriptive material on the range of all employers' experiences needed? Research design should follow these decisions; it should not drive them.

The major value of studying all termination employers is the ability to characterize employer benefits from JTPA. The gross impact approach does not allow precise estimates of the net value of such benefits, but it does provide a vehicle to describe them. The major value of studying participating employers is program development. Analysis can identify variants which enhance employer satisfaction and those which enhance the value of employer-based services to participants.

Multiple Employer Placements

Any analysis of participating employers must include all participating employers within the population to be analyzed. In the event that one participant is placed with more than one employer before program termination, both employers must be included. To identify the sample of participating employers only from the population of those who were the "final" employers, is to bias the employer sample by eliminating a group of placements which worked out especially poorly -- those which ended prematurely and were followed by transfer to further treatment.

This means that sample selection procedures must insure that these employers are included in their representative proportion. Where the sample is integrated with measurement of post-program performance standards, a supplementary sample of participating employers will be identified in order to enlarge the number of participating employers analyzed. Concomitantly, the records of former participants included in the original sample should be checked to identify earlier employer placements.

DESIGNING THE SAMPLE

Sample Size

For the most part, the discussion of sampling which appears in Chapter 4 covers employer sampling issues as well. Only one issue differentiates the two and must be discussed here: calculating the number who must be identified initially in order to produce a given sample of *completed* interviews differs for employers and participants.

The completion rate for employer surveys is likely to reach 90% or more. The major reason for nonresponse among participants is that they move and are not locatable. Employers move less often, and they encourage easy tracking when they do move. Therefore, aside from employers who go out of business, the only blocks to completing employer interviews is that a few will refuse and a few more will be busy or in the field so often that continuing to call back becomes inefficient.

This means that the initial sample of employers required to produce a target sample of completed interviews is *smaller than with participant samples*. For example, if we decide to aim for 400 completed interviews and expect a 90% completion rate, an initial sample of 444 will suffice. For a participant survey with 70% completion rate, the figure would be 571.

Another element is added if a sample of participants is selected at enrollment and a joint analysis of participants and employers in that sample is planned. In that case, only a portion of the participant sample involves placement employers, with other participants terminating unemployed. In such cases, the number initially selected into the sample should be enlarged by dividing by the probable proportion who will be employed at termination.

When the Employer Sample Should Be Identified

As with the participant follow-up, the sample of employers may be identified at termination or at the point of the participant's enrollment. For samples including only *termination* employers, employer names must be identified at termination.

When *participating* employers are being analyzed, contracts which end without a participant termination must be included in the sample. This means the sample may be identified at one of two junctures:

1. Employers could be identified during the process of ending contracts, such as at the point of final payment, or

2. A sample of participants could be identified at enrollment in employer based programs, with all the employers participating in their treatment constituting the employer sample.

Integrating Employer and Participant Samples

The combination of employer and participant data is recommended for any but the most basic descriptive analysis of employer benefits. If both employer and participant follow-up analyses are conducted, samples should overlap as much as possible. The validity of each in no way depends on the degree of overlap between the participant and employer samples; it depends on the representativeness of each sample. Therefore, additional analysis possibilities can be developed by maximizing sample overlap as much as possible without reducing representativeness.

That is done by using the participant sample to identify the employer sample. The sample of potential follow-ups, not the sample of completed participant interviews, should be used. Since these participants were selected at random, their employers are also a random sample. If the resulting sample of employers is too small, additional employers may be added using the sample selection methods discussed in Chapter 4. Even with complete overlap when samples are identified, some cases will not generate completed interviews for both participant and employer. However, high employer completion rates insure a substantial number of cases in which both types of data are available for joint analysis.

WHAT FOLLOW-UP PERIOD IS RECOMMENDED?

A single three-month follow-up survey should be adequate to measure employer benefits. However, extending that follow-up period in order to coordinate with participant follow-ups is acceptable. The longer follow-up period should increase non-completion only slightly among employers. Coordinating follow-ups, where possible, is the highest priority. Aside from that, three months appears to be a reasonable follow-up period, with the exception that participating employers who do not retain their participants at termination could be interviewed immediately following the end of their contracts.

Termination Employers

- Three months allows a sufficient time for most clearly unsuccessful jobs to have ended.
- Three months allows a sufficient time for all or much of a typical probationary period to have passed, so that the employer can report not only whether the participant is currently employed, but also how likely that employment is to continue in the future.
- After three months, the employer should be able to offer somewhat reliable estimates of the likelihood that the participant will be promoted, will receive a raise, or will receive increases in benefits packages in the forseeable future.
- Sufficient time has elapsed for employers to observe the participant's work. Employers will therefore be able to rate both how much post-JTPA training was required for the participant in question, and how well the participant is able to perform the work.
- In the case of participating employers, enough time has elapsed for solid evaluations of the worker to form in absence of whatever effect is created by salary reimbursement.
- The time elapsed since the hire or since the end of training is short enough to allow accurate recollection of the the hiring process and of the types of training provided.
- Longer follow-up periods are useful primarily for examining the erosion of program impact and the transfer of skill. Both of these require follow-up information from participants on employment patterns beyond the termination job. Relatively little is to be gained from an extended employer follow-up.

Participating Employers

Those participating employers who retain their participants at termination are included in the above discussion of termination employers. However, for those who do not retain their participants, little is to be gained from waiting three months for an interview.

Aside from fresh recall, the primary reason for accelerating follow-up interviews with employers who do not retain their participants arises when the follow-up interviews are being conducted by local program staff. Typically, staff will wish to talk with employers about placements which did not end in retention, either to reassure the employer, to decide whether to avoid the employer in the future, or to improve future client referral for that employer. The service questions staff will ask are essentially the same as those which would be asked during an employer follow-up, making it logical to merge the two. Indeed, service providers may wish to consider routinely administering the employer follow-up form to all employers as a non-threatening way to elicit some of the information they need for their review of the placement.

Defining the Beginning Point of the Three Month Follow-up Period

The point from which the three month follow-up is defined is clearer for employers than for participants. However, it differs for termination employers and for participating employers. For those who are termination employers, the employment start date also begins the three month clock. Since program termination typically occurs soon after employment, it provides an adequate approximation in these cases.

For participating employers who do not retain their participants at termination, the beginning of the post-program follow-up period should be calculated from the *end of the placement contract.* This is true whether interviews are to occur immediately or after a three month time lag. Starting the clock at that time is also administratively convenient, since these employers must be identified during the process of contract termination.

DATA COLLECTION METHODS

Some data collection issues are specifically related to measuring employer outcomes, and must be discussed here.

The Validity of Survey Data

Some **response bias** problems may be introduced by the fact that participating employers may wish to participate again. In particular, if they are being interviewed by local program officers, they may be less than candid about their costs and benefits from participation. The primary decision here may be whether local program officers should conduct employer surveys, an option which has considerable programatic benefits, or whether surveys should be confidential and conducted by a neutral party, an approach considerably more protected from response bias.

Viewed only from the standpoint of measurement validity, research efforts must be neutral. If the results of an employer analysis are to be disseminated publicly, both the employers responding to the survey and the research consumers must be assured of the neutrality of the the measurement and the analysis, and of the confidentiality of individual responses.

However, the programmatic advantages of conducting interviews with participating employers locally are great. Especially in the case of participating employers, managers may wish to integrate the program debriefing and quality review with employer data collection efforts. In such an event, two options are open:

1. The results of the analysis may be used for internal program planning but not presented publicly as valid research, or

2. The survey portion of these employer contacts should be accomplished through use of a written form personally handed to the employer (to insure high return rate) but returned separately, through the mail. If analysis is local, the return must be anonymous, which has the serious drawback that anonymity prevents employer data from being merged with MIS or participant data. If analysis is conducted by contract or at the state level, the employer survey can include the identification code of the participant in question, allowing data to be combined.

In either event, steps can be taken to reduce employer response bias. Employers can be assured that their data are to be used for analysis rather than to influence resource allocation (if this is true.) Questions can be posed in evaluatively neutral terms. Questions can ask for behavioral reports, which are less susceptible to bias. Loaded or threatening questions can be avoided.

Most importantly, employer fears that their answers might be used against them may not be evoked at all if questions are couched in terms of employer reports on specific participants. This approach evokes for the employer the orientation that only JTPA clients, and not employers, are being scrutinized by the research. Responses which might imply low service levels can be interpreted as specific to that one participant, rather than general to that employer. Indeed, it is true that individual reports are not a proper basis for establishing service levels. These surveys produce valid conclusions only over an aggregate of employers rather then individual by individual.

What Form of Survey Should Be Used?

- When employer data are to be collected by anyone other than local program staff who work directly with employers, telephone interviews should be used.
- With local data collection, the same is true except:
    - -- Where research design purity is compromised slightly in order to integrate employer interviews into the JTPA service delivery program, as discussed just above, a written form is more protected from bias.
    - -- If large employers with multiple JTPA referrals prefer, a written form can be completed routinely by each participant's direct supervisor.

Measuring Employer Training and Work Experience

When one goal of employer surveys is to perform differential impact analysis, characteristics of the employment establishment and of the selection and training of particular participants will be included among employer measures. Three levels of measurement specificity are encountered:

1. Measures characterizing the entire employment establishment, such as number of employees, industrial sector, or referral patterns established with JTPA.
2. Measures applying to any employee with the same job held by the JTPA participant, such as job complexity, qualifications required for that job, or training level of typical non-JTPA hires.
3. Measures applying specifically to each JTPA participant, such as the length of training received or the employer's ratings of that participant.

In cases where a single follow-up study with no integration into ongoing program activities is planned, selected measures of all three types may be included as part of the employer follow-up survey. Also, when employer follow-ups are being administered by the state, the research process may be simplified by integrating all data collection into a single follow-up instrument.

On the other hand, in cases where SDAs envision repeated local employer follow-ups, efficiency can be increased and nuisance to employers decreased by treating categories 1 and 2 as once-only measures analogous to those for service providers. Type 1 measures could be taken during an initial work-up with each new employer. Type 2 measures would be gathered once for each separate job title into which each employer accepted JTPA participants. Both these sets of information are presumably similar to information which should already be routinely elicited, but perhaps not formally recorded, before service providers decide to write a contract with an employer. Therefore, such measures are easily integrated into program operation where employer *or* participant follow-up analyses are envisioned.

SUMMARY OF RECOMMENDED EMPLOYER DESIGNS

1. Decisions to study terminating employers and/or participating employers, and to describe employer outcomes only or to analyze employer data in combination with participant data are policy decisions, which should be made prior to specific research design decisions.

2. Terminating employers are most easily identified by sampling participants who enter employment at termination.

3. If participating employers are analyzed, employer assignments that end in a transfer to other JTPA services must be identified at the close of each contract and included in the analysis.

4. Sample size considerations are identical to those for participants except that fewer employers must be sampled to reach a given goal of completed interviews because the completion rate will be higher.

5. If both employer and participant follow-ups are planned, the same sample should be used for both and the analysis integrated.

6. Where employer and participant data will be combined, the employer follow-up period should be set equal to that for participants. Otherwise, for termination employers, a three month follow-up is recommended. For participating employers who did not retain participants at termination, local follow-up can be accomplished immediately or after three months.

7. The follow-up period should be counted from the beginning date of unsubsidized employment or the end of a contract, whichever is earlier.

8. Employer follow-ups should be telephone interviews unless local programatic reasons indicate to the contrary.

9. Measures characterizing participating employers should be gathered once-only locally if possible, and through telephone surveys otherwise.

10. If results of an employer analysis are to be made public, care should be taken to assure employers that their individual interview responses will not be used to influence their future participation.

CHAPTER 6

SUMMARIES OF RECOMMENDED EMPLOYER AND PARTICIPANT DESIGNS

Chapter 6. Summaries of Recommended Employer and Participant Designs

This guide has dual objectives: to stimulate state and SDA planning by suggesting design and analysis options and to reduce planning effort by recommending particular designs over others. The last two chapters have raised a number of options and have offered numerous recommendations. However, they have not addressed the question of what overall, integrated designs are most advisable.

The bulk of this chapter is contained in one exhibit that presents abbreviated profiles of four particular designs. Two of these involve only participant or only employer follow-ups. The others integrate the two. These are offered to illustrate the strengths and weaknesses of various design choices, and also because each of these is able to maximize one or more criteria which might be used to select specific designs.

Following Exhibit 6.1 is a brief discussion of the two designs which integrate participant and employer follow-ups.

INTEGRATION WITH MEASUREMENT OF POST-PROGRAM PERFORMANCE STANDARDS

As this guide is being written, the advent of post-program performance standards requiring a 13 week follow-up of participants is very likely, although not yet definite. Some analysis designs are especially efficient and to be recommended in the absence of a required participant follow-up, but become less efficient given the form of measurement currently proposed for performance standards. In particular, collecting employer data is less efficient with participant contact required than if both data sources could be combined. However, these designs are included below because states or SDAs may wish to conduct program analysis separately from measurement of required performance standards. One example would be conducting a longer term follow-up beyond the three month requirement.

EXHIBIT 6.1 PROFILES OF FOUR MAJOR DESIGN ALTERNATIVES

I. Minimal Investment, Termination Employer Follow-Up Descriptions

Design Overview:

State or local sample of all termination employers, data from follow-up interviews only.

Major Descriptive Questions Addressed:

Descriptions of employer outcomes, including perceptions of JTPA and JTPA participants.

Measurement of participant retention of termination jobs is also possible, but is not recommended with this design.

Major Differential Impact Analysis Questions Addressed:

None. The addition of MIS and labor market data would allow minimal differential impact analyses. However, these analyses are better suited to other designs, and the simplicity of this design is one of its advantages.

Core Measures Included:

Employer perceptions of JTPA, of each participant they hired, of the participant's training, and of the participant's work.

Optional Measures:

Employer marketing information.

MIS Descriptions of participants placed with each employer.

Data Required From Agency Records:

Sample identification; MIS data if wished; program activity (e.g., employer-based or not).

Purposes Best Served:

Description of employer outcomes and perceptions. This design examines employer outcomes without reference to, or coordination with, participant measures.

Relationship to Required Post-Program Performance Standards:

No direct relationship.

Level Best Administered:

Can be state or local. Little information must be transferred between levels, reducing complexity of central coordination. Results are local, with little analytic power gained from combining many localities.

Sample Considerations:

Identify at participant termination. For participating employers, must refer to contract for best interview contact person.

High completion rate expected.

Bias Problems and Protection:

Participating employers who did not become placement employers are not included. This limits employer marketing value and prevents generalization to descriptions of all JTPA employers.

Measures characterizing former participants may not be generalized to all participants, since those who terminated without jobs are not included.

Investment Level:

Low. Interviews may be brief; interviews may be conducted throughout the day and with minimal tracking costs; little or no data merging; data analysis not complex.

Summary of Major Strengths and Weaknesses:

Provides partial information on several different topics of value at low cost, but does not offer full information on any. During a short follow-up, retention of the termination job is the most important participant outcome, but not the only one of interest. Termination employers can give an overview of employer perceptions, but important segments of employers are omitted.

II. Minimal State Oriented Participant Follow-Up

Design Overview:

Minimal follow-up interviews of participants terminated from all activities, conducted and administered at the state level and including program implementation measures, along with control variables.

Major Descriptive Questions Addressed:

What are participants' post-program labor market experiences, at follow-up and between termination and follow-up.

What are the variations in agency implementation patterns?

Major Differential Impact Analysis Questions Addressed:

How do SDAs differ with regard to performance on the most basic post- program outcomes, after adjusting for necessary control variables.

What agency implementation patterns are associated with greater post- program success?

Core Measures Included:

- Whatever range of participant outcomes the state wishes to measure.
- SDA and program activity membership variables.
- SDA once-only implementation variables.
- MIS and labor market control variables

Optional Measures:

- Individual selection and treatment variables reported retrospectively by participants.
- Subcontractor membership.

Data Required From Agency Records:

- Participant sample names and locator information.
- MIS data for participant characteristics and basic program activities.

Purposes Best Served:

This design is directed toward states' roles of rewarding high performers, providing technical assistance to poor performers, and adjusting performance expectations where dictated by service levels.

Relationship to Required Post-Program Performance Standards:

Ideally suited for integration. Analysis can be performed with only the required performance standards as outcomes or with additional outcome measures. Sample stratification needs are accomplished automatically by the demands of SDA performance standards measurement.

Level Best Administered:

The state. A sufficient number of SDAs and carefully standardized measures are central to the purpose of this research.

Sample Considerations:

Analysis goals require inclusion of as many SDAs as possible. If SDAs vary widely in size, the sample should be disproportionately stratified by SDA to equalize samples from each.

Bias Problems and Protection:

Nonresponse bias is a potential problem with participant interviews. Accurate locator information and tracking effort minimize the problem.

Selection bias threatens analysis, but protection is offered by control variables: client characteristics, agency selection measures, and, if funds allow, a limited range of individual selection and treatment measures included in follow-up interviews.

Investment Level:

Total cost is moderate. Marginal cost above and beyond cost required to measure post-program performance standards is low.

A state interview of service providers is required to make differential impact analysis most useful.

Some data merging and moderately complex analysis are involved.

Summary of Major Strengths and Weaknesses:

Measurement of individual treatment variations and of employer outcomes are sacrificed to concentrate the design on questions of most use to the state oversight role and to minimize marginal cost increase over that required by performance measurement. If the state includes enough SDAs for such an analysis to succeed, estimates are quite reliable and directly supportive of major state JTPA responsibilities.

III. Stepwise Employer/Participant Follow-Up

Design Overview:

Three rounds of interviews are combined to estimate the full range of participant outcomes plus outcomes for termination employers. These include:

- Interviews with a sample of participants who terminated unemployed.
- Interviews with a sample of termination employers.
- A second round of interviews with participants identified during employer interviews as having lost their termination jobs.

Major Descriptive Questions Addressed:

Nearly all descriptive questions asked by the termination employer-only and by participant-only designs can be asked with this one integrated design.

Major Differential Impact Analysis Questions Addressed:

These depend on the optional measures selected. The defining characteristic of this design is its efficient integration of participant and employer data. Which data elements are measured from each, and whether agency implementation variables are also measured, are open decisions within this design.

Core Measures Included:

- Employer outcome measures (perceptions of JTPA and of participants).
- Participant outcomes which can be reported by employers in cases where termination jobs are retained: employment, earnings, hours, etc.

Optional Measures:

- Employer marketing information.
- Agency level once-only implementation measures.
- MIS and labor market control variables, if differential impact analysis is planned.

Data Required From Agency Records:

- Employer and participant sample names and locator information.
- MIS data for participant background and basic program activities.

Purposes Best Served:

This design provides a maximal range of descriptive outcome data per completed interview. With slight exceptions, it is capable of addressing any of the goals stipulated in either of the two simpler designs above, depending on what optional measures are included.

Relationship to Required Post-Program Performance Standards:

Impossible to integrate because of employer-based measurement.

Level Best Administered:

State or local.

Sample Considerations:

The division of interviews into three subsamples necessitates recombining completed interviews, including weighting the three segments to re-establish their representative proportions.

Bias Problems and Protection:

Differential impact analysis is slightly less well protected from selection bias than with other designs because individual selection and treatment variants cannot be measured during employer surveys, and employer selection cannot be measured in participant surveys.

Possible non-response bias is lower for retention estimates than for turnover or late entry estimates.

Investment Level:

Moderate. Sample complexity and start-up costs for two surveys add to the cost of this design. However, ease of completion for employer surveys reduces costs somewhat.

Summary of Major Strengths and Weaknesses:

Except for its difficulty integrating with the format required of post- program performance standards, Consumer Reports might label this design a "best buy." For a marginal increase in data collection cost and sample complexity, one study covers both the full range of participant outcomes and also all employer outcomes which apply to termination employers.

IV. Local Emphasis Program Development Approach

Design Overview:

Participant and, if desired, termination employer follow-up surveys are combined with the full range of program variant measures and control variables, with the goal of developing the most thorough possible differential impact analysis of local level service delivery patterns. Any or all program activities may be analyzed, including in-depth analysis of all participating employers.

Major Descriptive Questions Addressed:

For this design, the range of descriptive outcomes measured is less important than the differential impact analysis potentials.

Basic participant and/or employer outcomes can be described.

Individual selection processes and treatment variants can also be profiled.

Major Differential Impact Analysis Questions Addressed:

Any of the differential impact analysis questions covered in this guide can be tested in this approach, with the exception that for samples including few different SDAs, SDAs may not be compared.

Core Measures Included:

Selected outcome measures, along with the full battery of program variant and control variables:

- Agency level once-only implementation measures, if an adequate number of service providers is included.
- Individual selection and treatment (services) measures.
- Once-only measures of participating employers and jobs.
- MIS and labor market measures.

Optional Measures:

Outcome measures may be core labor market experiences only, or may include the full range.

Employer marketing information may be included.

Data Required From Agency Records:

More burden is placed on agencies for data collection and transmission in this design than in others. Measures include:

- Sample identifiers, and MIS records of individual background.
- Once-only agency implementation measures.
- Individual treatment measures collected by service delivery staff.
- Once-only interviews with participating employers collected by service delivery staff (optional.)
- Labor market information.

Purposes Best Served:

The extensive commitment to close-in measurement and the ability to attack the problem of selection bias from many angles suits this design to a focus on local program development through in-depth differential impact analysis.

If a decision is made to concentrate service delivery development on one basic program activity or one target population, this approach is appropriate.

Relationship to Required Post-Program Performance Standards:

Performance standards may be measured in conjunction with the participant follow-up interviews.

Level Best Administered:

State or local administration is possible. If such extensive data collection efforts can be agreed upon by many SDAs throughout the state, the increased variability and sample size can improve the analysis. Local administration by one large SDA or more than one small SDA is also appropriate, and is easier to coordinate.

Sample Considerations:

This approach would benefit from identifying a sample upon eligibility or enrollment, rather than at termination, because it includes measures of individual treatment throughout JTPA. However, integration with measurement of post-program performance standards probably eliminates this option.

Bias Problems and Protection:

The wide range of measures available in this approach increases the likelihood that key selection processes will be measured, thus reducing bias in final estimates of program variant impact. If the analysis is local, variation of program variants may be restricted, which can truncate sources of bias or variables helping to identify bias.

Reliance on participant interviews requires guarding against non- response bias.

Investment Level:

High. This approach is measurement intensive, including not only follow-up surveys but also individual treatment records and, optionally, measures of participating employers. Much of this expense is expressed as higher work loads for staff rather than as a separate contract or salaries for separate interview staff.

A variant, most appropriate for state level data collection, is to eliminate agency measures of individual treatment process and to add some of selected treatment measures to the participant follow-up form.

Summary of Major Strengths and Weaknesses:

This approach is the most costly of all, in terms of service provider staff time. In exchange, it is the design most capable of developing specific recommendations for local program development and improvement. Further, it allows serious differential impact analysis and therefore program development, without demanding that the sample contain large numbers of SDAs. It also offers the benefit of integrating well into local service delivery and quality control activities. This means some of the cost expressed as staff time is actually augmenting service delivery.

INTEGRATED PARTICIPANT-EMPLOYER DESIGNS

Stepwise Employer/Participant Follow-up

One of the designs summarized above, option III which combines employer and participant follow-up interviews in stages, has not been directly discussed in prior chapters. Until the advent of post-program performance standards requiring contact with participants, this design was the most efficient of all, and highly recommended. It is therefore included here, despite its inability to integrate with required post-program measurement.

One aspect of this design requires discussion here. Three major groups must be interviewed, and their completed interviews combined in such a way that each is represented in the final sample proportion to its presence in the full population. These groups are:

1. Participants who terminated without jobs can be interviewed at the point of three month follow-up.

2. All terminating employers can be interviewed after three months. This will include cases where the termination job was retained and where it was not retained.

3. Participants who were employed at termination but no longer hold their termination jobs will be identified during employer interviews. The participants can in turn be interviewed using a slight variation on the participant survey used for those who terminated unemployed.

Adding these together provides a full set of participant outcome data along with the employer interviews. The only complexity is that each category above will generate a different interview completion rate. Therefore, completed samples must be weighted so as to accurately represent the proportions of each group in the original JTPA participant population. Group 1 will experience typical participant completion rates. Group 2 will produce higher employer completion rates. Groups 3 will have the lowest rate of all since their numbers are reduced during two passes through interview attempts. The weighting process which occurs after data have been completed rebalances these three groups, returning them to their original proportions while leaving the total sample size unchanged. The example below illustrates that process.

ILLUSTRATION OF WEIGHTING COMPLETED EMPLOYER AND PARTICIPANT INTERVIEWS

Let us say we identified 1000 participants at their enrollment throughout the year. Of that sample, 73% terminated employed. We attempted participant interviews with the other 27% and encountered a 68% completion rate. We attempted 730 employer interviews covering terminations with employment, completing 90%, or 657. We discover that 23% of those, 151 participants, are no longer with their termination employers. We attempt to interview these participants, with a 72% completion rate. The resulting sample and the required adjustments are shown below.

	Original		Completion				Adjusted	
Category	Number	(%)	Rate		Number	Weight	Number	(%)
Terminated unemployed:	270	(27)	68%	=	184	1.17	216	(27)
Terminated employed, retained: *	562	(56)	90%	=	506	.89	449	(56)
Terminated employed, not retained: **	168	(17)	90%*72%	=	109	1.24	134	(17)
Totals	1000				799		799	

---

* Employer interviews are sufficient for this category.

** Employer interviews must be followed by participant interviews for this category. Therefore two completion rates must be taken into account.

Local Emphasis Program Development Approach

Other things equal, greater diversity and flexibility of measurement improves differential impact analysis, by widening the scope of policy implications, by helping to guard against bias from confounded program variants, and by providing more tools to adjust for selection bias. Therefore, analyses oriented to intensive program development will benefit from including not only individual treatment variants recorded by agency staff, but also data from participating employers describing the nature of the position occupied by the participant and of the training provided by the employer. The same is true of data describing the nature of training in classroom settings.

SECTION III.

MEASUREMENT

CHAPTER 7

ISSUES UNDERLYING INTERPRETATIONS OF THE JTPA INTERVENTION

Chapter 7. Issues Underlying Interpretations of the JTPA Intervention

Differential impact analysis requires the selection of program variants to be measured and analyzed. One basis for that selection is deciding which areas of JTPA service delivery hold the greatest potential for program development. Identifying these areas, in turn, requires knowledge of what the realistic, currently implemented program variants are, and an interpretation of how outcomes are influenced by program implementation and treatment. That is, the measures should be "grounded" in actual local practices and concerns. The most useful state and SDA differential impact analyses will therefore be based on program managers' knowledge of what program variants are in place and will test managers' interpretations concerning what is working for them.

In addition, past research findings also suggest useful measures. One goal of this guide is to stimulate consideration of which JTPA outcomes and program variants are valuable to measure. Section III of this guide begins with this chapter on the *nature of the JTPA intervention* to stimulate thinking about the selection of variables. Later chapters recommend specific measures.

Appendix A offers a selective review of previous research in the job training field. Conclusions of that review are *summarized* here, keeping this discussion brief.

CONCEPTUALIZING THE NATURE AND LOCATION OF JTPA PROGRAM INTERVENTIONS

Implementation, Services, and Outcomes

Throughout this guide, the distinction is made between *implementation* and *service* variables.

Implementation does not affect outcomes directly. It affects outcomes *indirectly*, by providing mechanisms for the delivery of services so as to produce the outcomes demanded under JTPA and established as local program goals. Services themselves should be seen as the key to producing observed program impact. The causal chain is long -- from national legislation to the post-program labor market experience of each individual participant. However, in that chain, the causal center of gravity is low: local and close to the individual.

Thus, in selecting program variants for inclusion in differential impact analysis, we must focus on services or on implementation forms which affect the nature, distribution, and intensity of services. This leads in directions somewhat unusual in formally published quantitative job training research but commonplace for program managers. It recommends "close-in" measures of the nature of services and toward seeing the process of assigning services as the pivotal point in the treatment process under control of JTPA service providers.

Types of JTPA Intervention

There is clearly no single JTPA intervention. Indeed, the JTPA lists 24 services allowed, although without precise definition of each. These differ in nature, and their nature is in some cases ambiguous.

Exhibit 7.1 Examples of JTPA Interventions

- An unambiguous zero intervention: An employer firmly hires an individual, then sends that person to JTPA to check whether he or she happens to be eligible for an OJT wage subsidy.
- An unambiguously small intervention: job search assistance only.
- An intervention with ambiguous intervention regarding short term employment: basic education.
- An intervention of variable intensity, unambiguously based on providing incentive for an employer to hire an individual temporarily: WEX.
- An intervention ambiguous with regard to intensity and also with regard to definition as training or as employer incentive to hire: OJT.
- An intervention unambiguously training-based, and of variable intensity: CT.
- Intense interventions ambiguous in terms of their nature and reasons for working or not: multiple training-and-other-services approaches.

Although the list in Exhibit 7.1 begins with zero intervention and ends with intensive intervention, there is no obvious way to rank those in between. They represent different treatment types, rather than degrees of one type. Further, each type may include a wide range of costs, which may in turn substantially alter the impact likely to result. For analysis of differential impacts, we must identify tentative types of JTPA interventions and test program variants relevant to each. As a first approximation, we might define three major types of intervention:

A. *Motivational or informational assistance,*[1] as in JSA or pre-program workshops covering motivation, work habits, or orientation to the work world. There is no reason to expect great impact from such minor interventions. Both cost and impact are likely to be small.[2] They are therefore intended either to accompany other services, as in the case of pre-training workshops, or to assist relatively job-ready individuals, as in the case of JSA.

B. *Interventions based on training* have their effect through changing the individual in ways recognized by employers and by participants themselves. It leaves as problematic the process of acquiring and retaining employment. The credentialling aspect of the training opens new employment doors. Job retention is improved by skill acquisition during training. Long term effects are intended.

---

[1] To the extent that different types of employment barriers are addressed through informational and motivational assistance, these two represent different types. They are combined here to indicate their similarity: they aim to change the individual, preparatory to the individual locating employment, but they do not involve training or other extensive intervention.

[2] From the standpoint of investment in human capital, a low cost intervention can have a small impact and still be worthwhile. These interventions differ from others not by lacking worth, but by representing a different approach to improving employability.

C. *Intervention based on employer hiring incentives* provide training, experience, or a temporary override of personal or interpersonal problems preventing the participant from getting a job. These interventions have in common that they can reduce the employment process from two stages to one: most participants need only *retain* the job gained through the hiring incentive. This benefit persists as long as that particular job is held. It is less clear how often these interventions also provide training or experience which transfers to the aquisition or retention of other jobs.

When and Where in the JTPA Service Delivery System Do the Major Determinants of Program Outcomes Occur?

The importance of identifying different types of interventions is that each leads to somewhat different points of measurement focus. A brief discussion of the three types identified above will illustrate.

Type A. Motivational or Informational Assistance. These interventions are typically under the control of SDAs or their subcontractors, making them readily accessible to program development efforts. On the other hand, except for formal job club approaches, it is far from clear that these services have any impact on labor market outcomes, or how they should be altered to maximize their effectiveness. Instead, they tend often to be part of the pre-or post-program *process* which helps a participant "hang in there," much like support services. This role may be extremely important. However, it is not easily captured during analysis of program outcomes. In addition, these interventions tend to involve either the most or the least job-ready, making it complex to establish the nature of appropriate outcomes which should be expected. Much less previous work guides expectations for these services than for others.

This discussion holds several implications:

- If the funding for data collection beyond that required by performance standards is at issue, these programs may be first candidates for elimination from a gross impact analysis. The recent expansion of JSA as a sole service makes it valuable to analyze. However, the core outcomes measured for performance standards may represent a reasonable base for early analysis efforts for such low cost, low impact programs.

- When individuals who receive JSA only are included in analyses, they may be treated as a *minimal intervention* group, for purposes of comparison with other services. However, great care must also be taken to examine the degree to which they are a *minimal need* group. That is, one must beware of selection bias.

- Given the role interventions of this type play in introducing participants to job training or keeping them involved, it may be that the most appropriate analysis would test their impact on the intermediate outcome: *program completion.*

- Any analysis of how effective these services are must separately define cases where they occur in combination with other services and cases where they constitute the entire service.

- The high degree of control SDAs have over these services and their short duration make them ideal candidates for true experiments. For example, if there are too many JSA applicants to be accommodated at one time by the job club phone bank, some will have to wait one or two weeks. If groups to be served immediately and to wait were selected at random, and the untreated group's program entry were delayed somewhat, the two could be compared on short run outcomes, raising only minimal ethical questions.

Type B. The Training Intervention. This intervention is the most widely recognized and practiced of all, occurring society-wide in many different settings. Most, but not all, of these interventions are controlled by schools rather than by JTPA. JTPA agencies exercise control primarily through choices among referrals. This type of intervention can be powerful, but is incomplete. The intervention itself does not usually end in employment. Instead, it *changes individuals* as a means to produce the final outcome, employment. This gives JTPA three major points of control over the success of these interventions, each of which suggests particular sets of measures.

The first is quality control over the mix of potential referral destinations, schools and training courses.

Second is the selection of the correct match between participant characteristics, labor market characteristics, and program services.

Third is the post-training period when the intervention must be completed through the successful transition to work.

In addition, since this intervention unfolds over time, SDAs may play the process role of offering supports to help keep particular participants enrolled. Needs-based stipends are the prime example.

Some implications of this discussion follow:

- Since this intervention has its impact primarily through increasing individuals' skills, it is efficiently utilized only when employment is located in the training field, where the credentialling and skill transfer have impact. Employment outside that skill area constitutes program failure in almost the same degree as absence of employment, except where a general credential or personal redefinition occurs which transfers into higher employability in many areas.

- This intervention occurs more slowly than most others. It is most appropriate for individuals motivated to invest time in changing themselves (their skill levels and perhaps other aspects of themselves).

- The fact that the training intervention ends before the employment goal is reached raises the issue of combined interventions, especially those following CT with OJT or JSA.

- This same fact makes follow-up research especially valuable. Labor market demand may fluctuate more widely and slowly within any given field than in the total. Therefore, some participants may terminate without jobs in part because they are continuing to search for jobs in the field which utilizes their training. For such individuals, the lack of employment shortly after participation should be viewed not as a program failure, but as a program cost to participants, in the form of an extended job search that incurs foregone earnings.

- This intervention relies heavily on building into the individual the momentum to carry them through the job search period. This factor, taken with the importance of training-related employment, places *career orientation* at center stage in the causal chain leading to employment success in pure training programs.

Type C. Intervention Based On Employer Hiring Incentives. This intervention is based more on changing the behavior of employers than on changing participants. Employers are enticed to hire someone they would not normally hire. Participants are changed by the temporary receipt of earnings. To the extent that this intervention type is mixed with the training type, as in OJT, participants are also changed from the training they receive from the employer.

Employers' behavior may also be changed when they are induced to give participants more extensive training than that which would normally occur for new employees hired into the same job.

The true nature of the intervention occurring in each WEX or OJT hire is difficult to measure for three reasons.

First, this type of intervention is not as fully institutionalized as in classroom training. It is somewhat open to definition by each SDA and to negotiation between JTPA and employers.

Second, job training programs themselves have not reached consensus regarding whether these interventions should be based on the training model.

Third, most of the intervention is invisible to JTPA. Only the employer knows how much hiring practices are affected by the incentive and how much the OJT participant's training differs from that of non-OJT employees. Thus, even when the intervention is effective, it is difficult to know why, since the nature and extent of the treatment cannot easily be measured.

Also, for participants, the intervention is, in part, the outcome. This is especially true of OJT, where most employers understand that they are expected to retain participants unless they perform unacceptably. This means that most of the intervention under the influence of JTPA has already been completed when the contract is signed, often the very moment when the participant is enrolled in JTPA!

Some implications for measurement follow:

- So little is known about this type of intervention that the standardized measures of individual treatment suggested in this guide may provide particularly valuable description, even before the analysis of differential impacts is conducted.

- It is extremely valuable to include employers in the measurement process, and to interview them concerning each participant they hire.

- The outcome most uniquely sensitive to this type of intervention differs from outcomes for other interventions. It is *retention with the placement employer.* Other outcomes are of course mandated and valid. Indeed, increased earnings and long term employment are more fundamental to the JTPA mandate. However, this outcome follows most logically from the nature of the intervention, and appears to be most sensitive to factors under the control of program managers. There is also evidence that it is closely associated with increased earnings and other job quality, at least on the short run (Simpson, 1984a.)

- Service provider *quality control* over OJT and WEX placements represents the point of major JTPA leverage over this type of program activity, and is therefore important to measure.

- JTPA service providers have great control over the implementation of OJT. However, the need to entice employers into participation has led most to pass that control to employers. This relative balance, affected by policy toward marketing, may be a critical implementation aspect of the intervention.

Conclusion

In addition to orienting us toward particular locations in service delivery systems for the measurement of program variants, this discussion of different types of interventions involves an issue especially serious for the most common form of differential impact analysis: the comparison of basic program activities. The considerable differences among these interventions raises the question of how to interpret comparisons among them. They typically involve different *selection processes, mechanisms affecting participants, timing of those effects,* and *major outcomes.*

Chapter 3 argues that comparisons among major program activities are the hardest to estimate reliably. This chapter's discussion also points to difficulties, since these interventions differ in such basic ways. Such comparisons should be made. However, special attention must be paid to the issue of selection bias, and findings must be interpreted keeping in mind that different activities achieve their intervention at different points in the delivery system and with maximum impact on different outcomes.

PROGRAM IMPACT VERSUS EFFECTS OF INDIVIDUAL BACKGROUND CHARACTERISTICS AND LOCAL LABOR MARKET CHARACTERISTICS

Before deciding whether to invest in differential impact analysis, states and SDAs might wish to consider whether there is any room to identify program effects after individual participant characteristics and local labor market characteristics are accounted for. These control variables often show the most powerful impact among variables tested in published studies. (See Appendix A.) Previous research suggests the following conclusions:

- The power of individual and labor market characteristics, as indicated by their combined *variance explained* in training program outcomes, varies from over 40% to well under 10% in different studies.

- The amount of power these variables have to explain variation in any one study is greater when:

    -- The level of analysis is the SDA as opposed to the individual, and therefore confounds unmeasured program and geographic factors with SDA reports of average client characteristics.

    -- The outcome under study is earnings rather then employment, or employment rather than training-related employment.

    -- The data base is national and includes confounded effects of regional economies.

    -- The analysis includes few program variables, leaving these confounded with individual attributes in some cases.

- For state and SDA differential impact analysis, the effect of individual and geographical variables are minimized by all four factors listed above. Differential impact analysis:

    -- Operates with the individual as unit of analysis,

    -- Offers choice among outcome variables,

    -- Is based on state or local samples, and

    -- Can include any program variables that can be measured.

- Program selection mechanisms tend to *level* individual differences among participants, making such differences less powerful among program participants than among the general population.

- Some evidence exists that "creaming" may sometimes reduce program outcome levels, as well as sometimes improve them. Where placement jobs are of modest quality, the most capable participants may quit more often, while the least qualified participants, who have fewer labor market options, tend to retain whatever gain JTPA offered.

Conclusion

Differential impact of program variants can be effectively studied. However, it is necessary that individual background characteristics and other control variables such as the local labor market environment be included in differential impact analysis. Failure to include them will result in selection bias that will invalidate conclusions based on the analysis. However, the impact of these control variables is unlikely to account for all differences among program outcomes. It is therefore valuable for managers to invest in proper analysis in order to identify program variants which affect outcomes.

ARE OUTCOMES AT TERMINATION AND AT FOLLOW-UP PRODUCED IN THE SAME WAYS?

Once we enlarge the range of JTPA outcomes of interest by including post-program labor market experiences, we must consider whether the same factors determine success at follow-up as at program termination. The outcomes themselves remain unchanged. However, the selection of variables for inclusion in differential impact analysis of post-program outcomes depends on correctly conceptualizing the causal paths beyond, as well as prior to, termination. In large degree, of course, the factors which affect employment status at one time also affect it at a slightly later time. However, three differences are useful to discuss.

First, because participants' performance at termination is critical for service providers' financial survival, various methods are practiced by providers to maximize the value of specified performance measures. In varying degree, these practices mean that termination figures are biased by agency recording procedures. Follow-up measures are less susceptible to such manipulation and are therefore more directly comparable across agencies, making them preferable for program analysis.

Second, different interventions produce a different sequence of post- program outcomes. In particular, classroom training benefits tend to begin slowly and, up to a point, increase over time, while OJT benefits tend to peak either at referral or at program termination. (See Appendix D.) One implication of this difference is that agencies will experience some trade-off between termination figures and follow-up figures. This is especially the case in OJT, where counting marginal placements on the very day the contract ends can increase employment rate at termination but decrease retention rate.

Another implication is that follow-ups should include *both individuals placed and those not placed at termination.* While a common approach has been to ask only whether those employed at termination retain their jobs, any valid comparison across program activities must allow tests of delayed program effects (late hires), as well as tests of the decay of program effects (non-retention.)

Third, the constellation of factors which lead to success in securing a job is probably somewhat different than the factors leading to job retention after hire. For example: training credentials and effective job search behaviors presumably affect hiring more than later job retention; a stable record of previous employment or high satisfaction with one type of work may predict retention more than success at finding employment.

This distinction directs our attention to the pre-program diagnosis of employment barriers. Some individuals may appear "job-ready" because they find work easily. However, if they typically lose interest in work or create interpersonal difficulties during employment, they may not be appropriate for most JTPA interventions. Barriers to locating employment are dealt with more easily than barriers to retaining work, either through providing entry level training or through hiring incentives.

This issue also has implications for JTPA service to groups targeted as especially needy. The most job-ready have post-program options from sources other than program participation. If the position they gain through JTPA is not satisfactory, they may leave. Those with fewer options outside their JTPA jobs may retain even relatively low quality JTPA jobs. Thus, determinants of post-program outcomes differ depending on pre-program employment options.[3]

ISSUES HIGHLIGHTED IN JTPA IMPLEMENTATION STUDIES

Recent JTPA implementation studies have identified several issues which are either uniquely characteristic of JTPA or especially troublesome to early JTPA. (See Appendix A.) These may prove especially valuable for inclusion in differential impact analysis.

- The involvement and power of the private sector in general, and Private Industry Councils in particular are issues of importance.
- JTPA program benefits to employers have become a central topic of discussion as legitimate outcomes of JTPA programs.
- The central problem identified by implementation studies is the conflict between serving those most in need and producing high performance ratings through enrolling the cream of the eligible crop.
- Cost concerns appear to have led most SDAs toward short, and perhaps weak, interventions. Differential impact analysis of these compared to more in-depth interventions may aid our understanding of the trade-offs between individual needs, treatment intensity, and costs.
- The limitations placed on CT living stipends have produced a new generation of service delivery to examine. There appears to be more than adequate variation in level of CT enrollment and in average level of CT stipend to allow differential impact analysis of these issues.
- Centralized intake versus intake by subcontractors is also identified as a policy area of interest, with greater centralization under JTPA than under CETA.
- The nature of contracts and subcontracts has also changed under JTPA, with a large increase in fixed-price contracts. While this change appears primarily to represent movement to efficient, business-like forms, it may influence program outcomes as well.
- Finally, a new interest in analysis of internal program data seems to have arisen. Many states and SDAs express interest in increasing both the use they make of internal data and the development of new measures of program operation and performance.

SUMMARY

JTPA retains the CETA characteristic of umbrella legislation including several distinct interventions, each involving different amounts of training, occurring at different times in participants' JTPA careers, involving different relationships with employers, and bringing different influences to bear on probable post-program success. These differences introduce special difficulties into any analysis including more than one program activity, but also enhance the value of performing such analysis, to learn more about what makes each intervention operate successfully. These interventions have in common that the factors most influencing participant outcomes are those at the direct service delivery level and that JTPA agencies have their greatest leverage at the point where assignments are developed.

---

[3] The difficulty, of course, is that pre--program barriers to employment are so difficult to measure.

These discussions about how JTPA works were undertaken in order to provide direction to the selection of measures most likely to influence JTPA outcome success levels. Regarding outcomes themselves, we are led to emphasize *skill transfer* and to *broaden the range of outcomes* measured. Regarding influences on those outcomes, we are directed to measure different *referral avenues, agency quality control patterns, the match between participant needs and the intervention arranged,* and *specific characteristics of the intervention* which the participant experiences.

CHAPTER 8
PARTICIPANT MEASURES

Chapter 8. Participant Measures

This chapter discusses the types of measures which might be used by states and SDAs pursuing descriptive or differential impact analysis. Discussions of measures are grouped by source and use, as set forth in Exhibit 2.1. This guide also includes a specific survey instrument which could be used for participant follow-ups and a partial outline of agency treatment measures, based on the discussion in this chapter. (See Appendix E.)

Because differential impact analysis involves several distinct types of measures, this chapter is written in four separate sections. One covers outcome measures, and three discuss possible influences on those outcomes: agency implementation policies and practices, individual treatment received, and control variables.

PART I.

MEASURES OF PARTICIPANT OUTCOMES

A hierarchy of outcomes may be arranged according to the extent to which they are required for a meaningful analysis of JTPA. Some states or SDAs may wish to include only a minimal core set of measures, making the follow-up as brief and inexpensive as possible and limiting their analysis options accordingly. Others may wish to mount a more comprehensive analysis, once the decision is made to expend initial set-up costs. The marginal increase in cost from inclusion of all five types of measures discussed below is small, making it logical to measure all. Nevertheless, some measures add information without being necessary to the research effort. With this distinction in mind, Exhibit 8.1 displays outcomes in five categories, from highest (1) to lowest (6) priority.

EXHIBIT 8.1 PRIORITIZED PARTICIPANT OUTCOME MEASURES

1. **Required post-program performance standard standards.**

   - Employed during the 13th week after termination?
   - Earnings during the 13th week.
   - Number of weeks worked, during 13 week follow-up period.

2. **Other core measures explicit in the JTPA mandate.**

   - **Employment**, including:
     Hours per week employed at follow-up.
     Pre- to post-program change in hours per week and proportion weeks employed.

   - **Earnings**, including:
     Hourly wage rate at follow-up.
     Total earnings from termination to follow-up.
     Pre-program to post-program change in wages and earnings.

   - **Welfare dependency**, including:
     Whether receiving public assistance at follow-up.
     Monthly dollar amount of public assistance at follow-up.
     Total public assistance received between termination and follow-up.
     Pre-program to post-program change in public assistance received.

3. **Measures of skill transfer and utilization.**

   - Whether employment is in a training-related field.
   - Proportion of the work that utilizes skills from training.
   - For employer-based interventions, retention with that employer.

4. **Measures of job quality.**

   - Benefits (medical, retirement plans; paid vacations; sick leave).
   - Likelihood of layoffs.
   - Stability of hours worked.
   - Likelihood of promotion and/or raises.

5. **Measures characterizing those not employed or not retaining jobs held at termination.**

   - Why termination job was lost or left, if applicable.
   - Whether participant remains in the labor force, and if not, why not.

6. **Subjective orientations of participants.**

   - Intention to make use of the JTPA intervention (career orientation).
   - Personal evaluation of JTPA program services.
   - Personal evaluation of post-program job.
   - Personal comparison of post-program job with pre-program job.

The most basic outcomes focus on the **explicit JTPA mandate** that JTPA be considered an investment in individual lives -- an investment in human capital. As such, it should show returns in higher probability of employment, higher earnings, and lower dependence on public assistance. Using the survey method to collect data allows several components of employment and earnings to be specified. These same variables can also be measured at enrollment, covering the period prior to participation, making pre-post program comparisons possible.

Some of these core outcomes have been selected as JTPA performance standards. By definition, these are first priority to three month follow-up efforts because they are required. In addition, for reasons of comparability, they should be treated as highest priority for analysis which covers longer follow-up periods. Other measures most explicit in the legislation are listed as second priority to follow-up efforts, and should be included whenever analysts wish to measure more than performance standards.

Of third level importance are measures indicating **skill transfer and utilization**. This outcome is not as explicit in the legislation, although it is implied. It represents the most direct impact of those interventions based on training, and is especially sensitive to program variants, making this outcome particularly useful to managers who wish to develop their programs based on differential impact analyses. (See Appendix D.)

A fourth level of importance may be assigned to measures of **job quality.** In addition to wages, various intangible benefits from employment and indirect forms of income such as medical benefits, are important aspects of job quality. A prime indicator of probable long range employment success is whether the overall quality of each job places it into the category sometimes characterized as the "primary labor market" or into the "secondary labor market" (Doeringer and Piore, 1971; Vermeulen and Hudson-Wilson, 1981). Primary labor market jobs are relatively stable, include gradually improving income and benefit levels, are usually full- time, include the possibility of promotion, include fringe benefits and are, in general, the type of job which can reasonably become a career.

The fifth priority type of outcome measures **aspects of unemployment**. Interpretations of program effects may benefit from knowing the reasons why participants lose their jobs or leave the labor force, and the outlook for future employment among those currently without work.

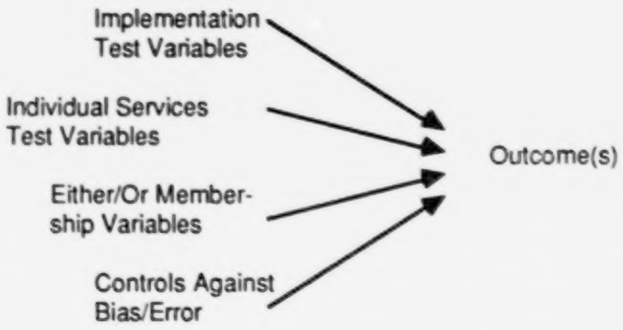
Finally, the lowest priority outcomes are **personal attitudes and orientations of participants**. These occupy sixth level because their meaning is less clear, they are less reliably measured than other outcomes, and they have been excluded from most job training evaluation studies. (See Appendix A.) However, they can offer valuable information to JTPA program operators and their measurement is inexpensive once participants are being interviewed. They also give the participant an opportunity to express his or her feelings, which can help cement rapport if future interviews are planned. Therefore, analysts who plan full data collection at six months, following measurement of performance standards at three months, may wish to include satisfaction measures in the three month interviews.

The participant interview instrument in this guide includes measures of each type listed above. (See Appendix E.) Where the decision is made to invest in shorter surveys, the priority listing above represents a set of recommendations for which measures are most important to retain and which can be eliminated with less serious loss of information.

ORGANIZING A LIMITED SET OF POTENTIAL INFLUENCES ON OUTCOMES

Differential impact analysis tests the impact of program variants and control variables on post-program outcomes. Basic categories are shown in Exhibit 8.2, below.

EXHIBIT 8.2 BASIC ELEMENTS OF DIFFERENTIAL IMPACT EQUATIONS


Implementation Test Variables
Individual Services Test Variables
Either/Or Membership Variables
Controls Against Bias/Error
Outcome(s)

Two of these sets of measures, the *test variables*, are identified by choice, because the analyst hopes to learn whether particular implementation forms or particular individual services enhance program outcomes. A third set, *control variables*, is required in order to insure that the effects estimated for test variables are as accurate as possible. The final set, *either/or membership* variables, indicate concrete categories such as a particular SDA or a particular program activity. These are useful both as test variables and as control variables.

WHEN STATES OR SDAS IDENTIFY IMPLEMENTATION OR TREATMENT VARIABLES FOR ANALYSIS, THE FOLLOWING CRITERIA ARE USEFUL

1. Can the variable be measured reliably?
2. Is there reason to believe it varies across individuals or service providers?
3. Is there reason to believe it represents a non-trivial program impact?
4. Does the variable measure a program variant under the control of program managers? That is, is it a policy variable?
5. Are program managers open to changing the program variant to be measured?
6. Given cost and time constraints, can the measure be integrated into a data collection scheme?

Criteria 1 and 6 in the above list are essentially technical, and may be answered without much knowledge of JTPA. However, criteria 2-5 require knowledge of the state or local JTPA service delivery system, making input from program managers critical to successful analysis.

The remainder of this chapter discusses measures of each type listed in Exhibit 8.2. For some measures -- outcomes, individual treatment variants, and control variables -- this guide suggests specific measures. (See Appendix E.) These are intended to be usable by organizations wishing to minimize start up costs and also to act as a stimulus to others wishing to develop their own instruments.

In the case of program implementation variants, measures are discussed, but no specific measurement instrument is provided. Variation among states and SDAs is too great, and policy development interests too varied, for this guide to develop specific implementation measures. These must be locally grounded. In addition, JTPA implementation studies and some excellent analyses of CETA implementation also suggest measures. (See Snedeker and Snedeker, 1973, Levitan and Mangum, 1981; Franklin and Ripley, 1984.)

PART II.

MEASURING PROGRAM IMPLEMENTATION VARIANTS

MEASUREMENT SOURCES

Program implementation variables measure aspects of the organization put in place to provide JTPA services. They characterize the entire organization -- its structure, policies, and practices -- rather than any one participant's treatment.

Most program implementation variants are best measured through surveys of service providing organizations. Data for each service provider can then be attached to computerized data files of all participants who enrolled through that provider. Depending on the nature of the measures, agency directors may be able to answer reliably, or agency staff may need to compare notes or to consult records in order to characterize typical practices accurately.

With easy-to-answer questions, phone surveys may be used. However, for more demanding measures, a written survey which allows time for data gathering is preferable. Our recommendation is for a written survey of each service providing organization, with a backup telephone contact person who can clarify questions as they arise.

One problem with this type of measurement is that the agency may *intend* one form of implementation but actually carry out another, making agency self-descriptions inaccurate. This can be partially remedied by a second form of implementation measurement: *aggregated agency characteristics*. Aggregated variables are measured with the individual participant as unit and then summed, percentaged, or averaged across all participants within each agency. (See Appendix B.)

For example, agencies could report whether their policies emphasize training women for non-traditional occupations. The aggregated form of measurement for this same issue begins by constructing the individual level variable. Training fields which are non-traditional for women are identified, and the training field of each female participant in the sample is coded as traditional or non-traditional. That individual variable is then aggregated for each agency, producing an agency level variable, *percent of female participants in non-traditional fields*. These measures are easily constructed by computer. However, the approach is limited to variables that can be measured at the individual level.

SUGGESTED MEASURES OF PROGRAM IMPLEMENTATION VARIANTS

Program implementation variants can be divided into *basic organizational components* such as forms of contracting and staffing, and the *service delivery framework* within which intake, referral, treatment, support services, and program exit occur. The latter are more likely than the former to influence program outcomes, because they affect the nature of services provided and the selection process through which individuals are assigned to treatments. Yet basic organizational components are easily observed, and may therefore arise as alternative explanations after the research is completed. ("Yes, but the differences you found only show up because this SDA has an excellent PIC, and that one is very small, and this one....") Their measurement provides inexpensive protection. If, in addition, any is found to affect outcomes, so much the better.

Basic Organizational Composition

Forms of Contracting. JTPA implementation studies suggest some evidence that use of Requests For Proposals (RFPs) as part of the SDA subcontracting process improves performance. Similarly, managers face the question of whether intake and other services should be centrally delivered through the SDA or subcontracted, and whether classroom training courses should occur in-house or through referral. For both service provider contracts and trainer referral contracts, use of fixed price, performance based contracts may be contrasted with other approaches to contracting.

Staff Qualifications. Franklin and Ripley (1984) argue strongly that staff qualifications represent a key to success, although they are not specific about what constitutes good qualifications. As a guide to staff training or hiring, analysts might test a set of staff qualification measures used by managers for hiring and promotion.

Staff Turnover. One might assume that staff stability (low turnover) would predict success, although during the late CETA era, one study found the reverse to be true (Simpson, 1984a). This may have been specific to that time period, when the definition of job training was changing rapidly, or it may be that staff burnout is a special problem.

Staff Workload and Division of Labor. Client : staff ratio, overall or within each program activity area, may relate primarily to service level,and may also affect outcomes. Similarly, the division of staff between direct service, administration, clerical support, and development work is valuable to describe, since it holds implications for agency ability to grow or stabilize in the future and may affect current outcomes.

Service Provider History. The age of service providers, how much their services have changed over time, their relations with the private and public sectors, their rate of growth or decline, and the like, may be useful to identify, although JTPA implementation studies suggest little differential impact on program outcomes.

Size. The size of SDAs or subcontractors (amount of grant, number of participants, number of staff) may also be included as control variables. While size is seldom open to revision on the basis of research, it is so obvious a control variable that differential impact research which fails to measure it may be discounted by consumers.

This list of basic organizational components rapidly becomes dangerously long. As discussed in Chapter 3, the number of implementation variables which can be included in any one equation explaining outcomes is relatively small. However, that problem is diminished, ironically, by the fact that few of these variables are likely to have noticeable impact on post- program outcomes. Therefore, most can be eliminated from the analysis one at a time, at the descriptive level of analysis.

Service Delivery Framework

Explicit Selection Processes. Agencies differ with regard to their selection processes and some of these differences can be measured. Agency policy may emphasize enrolling the most job-ready, those with greatest need, or those whom the program is most likely to benefit. Agency selection policy may treat different activities similarly or may reserve some, such as JSA or short OJT, primarily for those who are relatively easy to serve.

These issues of purposeful agency selection are critical to measure, both as a service quality issue and also because *selection bias can be partially addressed with such measures.* Measuring them *validly* can be difficult, but should be possible.

In addition, other factors besides explicit policy may influence agency selection. One such measure, available in only some cases, is the proportion of all eligible persons who are enrolled. The larger the pool to select from, the greater the likelihood of selection bias relative to the legally eligible population as a whole. This is true, however, only after the proportion of enrollments generated through employer referrals is accounted for. These enrollments show one enrollment for each individual found eligible, yet they involve high potential for selection bias because of their pre-selection by employers.

Intake Procedures. Procedures used during intake for selection, diagnosis, information giving, and counseling may differ in intensity and type. They may affect how well the agency treatments match the abilities and needs of each participant and labor market needs. They may also act as indicators of the agency selection process.

SUGGESTED TYPES OF INTAKE MEASURES

- Agency policy regarding intake may dictate full diagnosis by the agency, participant efforts to diagnose their own goals and abilities, or minimal diagnosis with assignments based on availability.
- Initial intake may be conducted primarily in group settings or individually.
- The tools available during intake diagnosis may vary. Diagnostic tests or workshops providing orientation to training or work may or may not be available. Staff may or may not make career counseling available.
- How often the available tools are used may vary. The number enrolled in workshops may vary, agency policy may be to test many or to reserve testing for the few cases where the expense can be justified, and the like.
- Policies regarding screening criteria vary. Different groups may be targeted as most important to serve; employer referrals may be automatically defined as suited for enrollment or may be reviewed; individuals with particular traits may be disproportionately referred to particular activities, and the like.

Quality Control Over Referral and Program Activity Mix. Service providers may exercise strict control over the development of participant assignments, including rigorous quality screening over which schools, agencies, or employers may be involved in treatment. At the other extreme, agencies may take the *laissez faire* approach, exercising as little control over the process as possible, choosing instead to facilitate whatever potential assignments arise. This issue promises to be one of the most valuable areas for agency level measurement, because referral represents the pivotal point of agency influence over treatment.

AGENCY CONTROL ISSUES

- What is the agency policy regarding employer referrals? Are they encouraged, giving control to employers, or carefully reviewed and screened, retaining agency control? Are they accepted from employers the agency has not previously screened?

- What proportion of placements with employers are initiated by employers?

- At the opposite extreme, does the agency encourage "open contract" type referral arrangements whereby employers agree to fill all openings for certain job titles by choosing among a set of eligibles sent for review by JTPA? What proportion of placements follow this path?

- Are participating employers reviewed for probable quality of job experience or training? Do reviews have teeth, or do agencies feel the need to arrange contracts with any employer willing to participate?

- Are participants encouraged to develop their own OJT placements, using agency materials confirming their eligibility for wage reimbursement? What proportion of OJT assignments are developed in this way?

- Does the agency formally review classroom trainers? How much impact do such reviews have over the mix of assignments?

The second half of the quality control issue is *how quality is defined.* Among those agencies which perform explicit quality control reviews over potential referrals or placements, what criteria are used?

QUALITY CONTROL CRITERIA

- Is previous JTPA placement or retention track record assessed, if applicable?

- **In the case of employer-based interventions**, do reviews determine:

  -- How stable the employer is?

  -- What the typical non-JTPA turnover rates and wage rates are?

  -- Whether each job meets a specified minimum level of skill?

  -- That cost-effective training will occur, if the intervention stipulates training?

  -- That the employer is capable of supervising constructively?

- **In the case of schools or community based trainers**, do reviews determine:

  -- Whether they are able to handle disadvantaged students?

  -- Whether they provide realistic placement assistance for graduates?

  -- Whether their credentialling is credible among employers?

Policy Toward Ancillary Support Services. For the most part, services ancillary to the primary treatment are aimed toward allowing the participant to complete a program or enhancing the quality of the training process. States or SDAs wishing to include provision of support services among the program variants they measure might therefore wish to analyze whether they influence the program *completion* rate, and how completion, in turn, affects post-program employment outcomes. Such a test is especially relevant to JTPA in the case of needs-based payments offered during classroom training programs.

Measures could include policy encouraging or discouraging use of support services; the proportion of participants receiving needs-based payments,along with their average level; and the proportion receiving other support services. All but the first of these are aggregate measures derived from measures of individual treatment experiences.

Exit Practices. The final element of treatment is the set of program completion and job search options which are implemented. In one case, JSA assignments, these constitute the entire program. There is clear experimental evidence that job clubs are effective placement mechanisms (Azrin, 1978), but we have little information about the retention rates they produce in less controlled field settings.

Exit practices are especially important to measure at the agency level. Individual measures of job search assistance suffer from the compensation problem: those least able to locate jobs on their own are most likely to receive job search assistance, creating the appearance that post-program employment is *negatively* correlated with receipt of job search assistance. For analysis of individual level treatment, accurate measurement of individual need is therefore required before the impact of JSA can be estimated. However, need is difficult to measure accurately. Agency level measures of the *average availability of assistance* can help fill that gap, since they are not affected by the compensation problem.

AGENCY EXIT PRACTICES

- Availability of job clubs and of job search workshops, and the proportion of participants aside from JSA-only clients, who use them.
- Proportion of agency staff time devoted to job development and referral activities for post-program placement.
- Proportion of trainers who include formal job placement assistance as part of their program.
- Whether placement is centralized or handled by subcontractors, and whether it is handled by the same individuals who handled intake, or by those handling training.
- Average agency funds spent on placement, by type of program activity.

PART III.

MEASURING INDIVIDUAL TREATMENT VARIANTS

Each participant's treatment consists of a set of services. Some of these are shared by all who enroll through a particular service provider or who participate in a particular program, while others vary by individual. Thus, the nature and intensity of services received by different participants vary widely *within* agencies as well as between them. These differences among individual experiences require individual level measurement.

MAJOR ADVANTAGES OF INDIVIDUAL LEVEL MEASURES:

- They tie program services to outcomes for the same specific individuals, offering precise analysis of the degree of association between the two.
- They also tie specific services and outcomes to specific individual background characteristics, providing direct tests of control variables.
- Normally, they vary more widely than agency level measures, strengthening statistical tests.
- Normally, they suffer less colinearity with other test variables than agency level measures, strengthening statistical tests.

MEASUREMENT SOURCES

Some individuals' treatment experiences are normally recorded as part of agency MIS files. Others may be:

1. Recorded by agency staff, as the treatments occur, or
2. Included in participant follow-up surveys, measured through participants' recall of the services they received.

The preferable form of measurement is agency recording. Agency staff can record services as they occur, avoiding participant recall errors at follow-up. Staff know the range of service options and are therefore more able than participants to identify which services are administered and when they occur. This makes measurement much more accurate and also allows a wider range of variables to be measured than would be possible through participant follow-ups.

Agency measurement is also less expensive, as measured in terms of dollars, since it avoids telephone interview time. It is, however, more expensive in terms of staff time. Many, of these measures are available through an effective MIS. However, many other desirable measures require additional data collection efforts.

Gathering such information through participant surveys also has certain advantages:

1. Less lead time is required. Where samples are identified at termination, this is the only option available.

2. State or multi-SDA analyses may find it impossible to coordinate data collection by large numbers of local direct service personnel, making measurement at follow-up the only viable option.

This guide provides a partially constructed agency treatment record. (See Appendix E.) It is left partial because completion requires knowledge of local service delivery options and requires knowing how much information agency staff members are willing to collect.

SUGGESTED MEASURES OF INDIVIDUAL TREATMENT VARIANTS

Screening, Selection, and Intake Services

Accurately estimating the impact of intake on program outcomes requires both measures of intake experiences and also measures of participants' need for intake assistance. Since intake intensity should to some extent be compensatory, good intake will tend to equalize the chances for success of those with greater and lesser need. The best approach to this problem is careful measurement of each participant's need upon application to JTPA and of the intake services they receive. Inclusion of both types of measures allows analysts to test them together in the same equation, so that the effects of both can emerge.[1]

Measuring Need for Intake Services. The problem is measurement. We have available no precise measures of need, and some aspects of need, such as emotional need, are unmeasurable.[2] We can approach this problem from three measurement perspectives:

1. Certain *objective individual characteristics* included in MIS files or measureable at enrollment are likely to be *correlated with need* for extensive intake. These include educational background and previous work history. Such correlates are only partial indicators of need for intake, but they are readily available and can be objectively measured.

2. *Participants' perception of their own need* for intake (e.g., statements that they know little about training options) are also presumably correlated with objective level of need.

3. *Agency staff judgements of need* for various intake services can also be recorded by staff. These judgements are correlated with objective level of need, and also represent the staff perceptions which directly influence the amount and types of intake services each participant receives.

The first of these approaches is subsumed under measurement of control variables, discussed later in this chapter. To construct measures using the other two approaches, analysts might identify areas of intake need such as those listed below. Agency staff members or participants could then rate each area on some standard scale, allowing the measure to be used during analysis. This approach has been developed most thoroughly in the area of youth competencies. (See Snedeker, 1986.)

---

[1] If compensation were perfect, colinearity would prevent estimates. Experience suggests this is unlikely to be a difficulty (Simpson, 1982; 1984a).

[2] In addition, these factors are distant from eventual program outcomes, so that their impacts are small. Nevertheless, they are valuable to quantify and they must be included in any effort to assess the impact of intake services.

POSSIBLE AREAS OF INTAKE DIAGNOSIS:

- Level of awareness of training options.
- Level of awareness of appropriate employment options.
- Level of experience or skill at applying for jobs.
- Participant's reported desire for a long term career change, versus for immediate employment.
- Participant's request for a specific placement or for employment in a specific field.

Except for the last measure listed above, program officers might score participants using the scale below. Scale values would represent:

1. Extreme need. Assistance required.
2. Substantial need. Assistance helpful.
3. Little or no need in this area.

Measuring Intake Services. Intake measures can include both the nature and intensity of services in the same areas for which need is measured.

MEASURES OF INTAKE SERVICES EXPERIENCED

- Time in initial group intake and/or initial individual intake.
- Time spent in workshops (zero hours indicates no workshop assigned). This could be measured separately for each type of workshop available. For example:
    - -- Orientation to training options.
    - -- Orientation to the work world.
    - -- Job search assistance.
- Time spent in individual counseling and information giving. Time could be broken into categories, as is done with workshops.
- Group and individual intake time could also be totalled, indicating overall intensity of the intake process for that individual.
- Specific intake experiences such as vocational testing or English language testing.

Another approach to intake measurement is to identify *separate paths taken by individuals enroute to program enrollment*. We sometimes err in supposing that there is a single phenomenon we may call "JTPA intake," which varies in intensity but not in its basic nature. The formally designated intake process is clear and widely acknowledged: an individual becomes aware of need, seeks assistance from JTPA, is found eligible, undergoes enough intake measurement for a JTPA staff person to offer guidance, receives the information and guidance, is assigned to a program and enrolled, and finally, enters the program. However, many do not follow this route.

Variations in intake route are important to measure for several reasons.

- They are measures of patterned differences in intake procedures which may affect outcomes.
- They are important as control variables to explain and adjust for intake selection.
- They involve service quality questions because apparent outcomes sometimes occur after no JTPA services except minimal record keeping.
- Finally, it is likely that simply describing intake in this way will provide useful information to program managers. For example, in our Washington OJT study every head of over 30 service agencies underestimated the proportion of their OJT contracts that were initiated by employers.

EXAMPLES OF INTAKE ROUTE VARIATIONS

- An individual is enrolled in school and comes to JTPA for tuition assistance, is found eligible, and is referred back into that same program.
- An employer locates a desirable employee during job interviews and sends that person to JTPA. He or she is found eligible and placed in OJT with that employer.
- An individual applies to JTPA, is found eligible, and is referred to the OJT program. He or she is given materials describing OJT to employers and sent out to search for a placement. Upon finding one, he or she is enrolled with that employer in OJT.
- After enrollment but before services begin, a job opportunity arises for a participant, who is then terminated and listed as having entered employment.

Delay Between Eligibility and Enrollment. The time lag between eligibility and enrollment may be a component of selection bias. Up to some point, delays tend to weed out the least motivated. However, long delays probably discourage those most qualified and motivated, too. This measure becomes interpretable only when analyzed together with the proportion of participants who enrolled after being referred by employers. These individuals are essentially found eligible and enrolled in one sitting. This prevents any delay, but also represents the possibility of marked selection bias because of their prior selection by employers.

Referral to Basic Program Activities

Clearly, one measure of individual treatment must be the basic program activity or activities in which each individual enrolls. These are normally available through MIS files, although information on multiple activities and sequencing will often require additional data collection effort.

Three other factors related to referral are also important to measure for each individual. First, *two types of multiple services must be separated:*

1. Multiple *sequenced activities planned in advance*, such as an orientation workshop followed by CT, followed by OJT in the same skill area.
2. *"Second chance" activities* assigned to individuals who failed to utilize their first service successfully.

Second chance assignments may or may not represent a good way to serve particular participants. What is clear is that those who failed once are more likely to fail the second time also. It is therefore important that this type of multiple activity program be analyzed separately from planned sequences.

The second referral issue important to measure is the *nature of each participant's referral to basic program activities.* As discussed under implementation measures, the specific referral or placement may be identified or initiated by schools, employers, JTPA agencies, or participants. Which of these applies in each individual case is important to measure, both for purposes of testing the impact of each and for protection against possible bias.

The third referral issue involves the *match between participant need and treatment*. Just as measuring competencies was suggested as part of analyzing intake, measuring barriers to employment is integral to assessing the impact of program activities.

.POSSIBLE EMPLOYMENT BARRIERS INFLUENCING ACTIVITY ASSIGNMENT:

- Level of basic literacy, required for some JTPA program activities.
- Level of proficiency with the English language, where an issue.
- Level of difficulty getting new jobs (as evidenced by pre-program record and/or by presentation of self during the JTPA interview).
- Level of difficulty retaining work (as evidenced by previous job history).
- Absence of previous work experience.
- Absence of previous training.
- Reasons why previous skill or experience no longer generate successful employment (where appropriate.)

Substantial differences are likely among participants selected into each basic program activity, some of which will bias estimates of each activity's impact. The inclusion of diagnostic variables such as those listed above can help identify and adjust for the effects of such differences.

Treatment Intensity and Completion

In addition to the type of program activity, the length and intensity of the activity should be measured. Further, whether participants complete their programs must be measured. Measuring program length alone would confuse the completion of short programs with the noncompletion of longer programs, making conclusions uninterpretable.

Length of (full-time equivalent) training and program completion are readily measured. Unfortunately, no precise measures of training intensity per time period exist. However, partial indicators of intensity can be measured once the intended nature of the intervention is identified.

POSSIBLE MEASURES OF THE LENGTH AND INTENSITY OF PROGRAM ACTIVITIES

- How long is the scheduled intervention?
- Is the scheduled intervention sufficient for a credential in fields where one exists, or to be competitive in a job market without credentials?
- What proportion of the scheduled intervention is completed?
- If less than 100% is completed, why did the participant leave?
  - -- Better employment opportunities.
  - -- Changes in life plans or situation.
  - -- Failure for personal or performance reasons.
- If the intervention is intended to provide training:
  - -- How many hours of training are scheduled?
  - -- How many hours involve relatively formal instruction?
- If the intervention involves employment experience:
  - -- How complex is the job, as reported by employers?
  - -- Does it also include training beyond minimum introduction?

Trainer Characteristics

Although factors such as trainer's methods or organizational forms can seldom be changed by JTPA, the knowledge of which types of trainers most effectively produce the outcomes desired by JTPA can improve JTPA quality control and referral decisions. In addition, information on effective training approaches may be of interest to schools and especially to employers, since relatively little is known about how to perform OJT effectively.

Some measures describing trainers can be gained from participant follow-up surveys. However, the most reliable measurement source is the trainers, schools or employers, themselves. Even where other sources of data are available, such as the *Dictionary of Occupational Titles,* job and training descriptions, direct descriptions of each individual training situation are far more useful (Simpson, 1984a). This guide includes a sample set of questions for participating employers. (See Appendix E.)

Classroom Training. The easiest measures of classroom trainers are also quite powerful because they are associated with cost, availability, public recognition, and probable effectiveness with various types of participants. These are *typologies of trainers*. For example:

- Is training:
    1. In-house,
    2. Referred to courses enrolling primarily JTPA participants,
    3. Referred to courses not limited to JTPA, but enrolling primarily individuals receiving other public subsidy, or
    4. Mainstream individual referrals?
- If not in-house, is the trainer:
    1. Public,
    2. Proprietary, or
    3. A community based organization?
- Is the trainer:
    1. Multi-purpose,
    2. Vocational only, with many programs, or
    3. Vocational, in only one field?

In addition, trainers vary in size, mix between experiential ("hands-on") and formal learning, inclusion of internships, and a great range of other characteristics (Simpson, 1982). Given the emphasis on explicit coordination between JTPA and vocational education systems, states or SDAs doing differential impact analysis might wish to develop particular measures of classroom trainer characteristics in collaboration with vocational educators.

Employer-Based Treatment. In the cases of OJT and WEX a wide range of measures is once again available. In this case, PIC members may play a useful role in the selection of particular measures. Relatively little widely disseminated research has been done on which to base recommendations for selection of measures. Our Washington state OJT research identifies some characteristics of the trainer and of the OJT position (Simpson, 1984a).

SELECTED CHARACTERISTICS OF EMPLOYER-TRAINERS

- The employer's growth rate.
- The typical turnover rate.
- The industrial sector, and whether public or private.
- The quality and complexity of the job.
- The amount of training offered.
- The use of relatively formal training methods.
- Whether skills gained from training apply to a wide range of jobs, or are "firm specific?"

Ancillary Support Services

The same ancillary services discussed under implementation measures can be measured as individual treatment variants. Except for the issue of stipends offered during classroom training, there appears to be no cogent reason to detail specific support services. However, the total amount expended per person can be recorded at little cost.

The most important service to study under JTPA, needs-based payments included with CT enrollment, may affect participants in two ways, which should be separated if possible. First, receiving a stipend may make the participant more dependent on JTPA, thereby encouraging program completion. Second, income from any source has the potential to affect life stability, personal stress, and other factors which can in turn influence post-program labor market experiences. Therefore, a precise analysis of the impact of stipend *per se* requires measurement of total income during training as well as income from stipends.

Program Cost

The primary *marginal cost* for each JTPA participant is the direct cost of training. Although *total* program cost figures are typically recorded in systems separate from MIS and involve calculation time lags before becoming available, marginal training costs for each participant are usually available through contracts with trainers or employers. (See Zornitsky, et al., 1985.) This means that any SDA wishing to perform differential impact analysis could measure the bulk of program costs attached to each specific participant. These could be analyzed along with other aspects of the program, to test improvement in program performance relative to individual treatment costs.

There are limitations on ability to interpret findings regarding costs. Testing cost as a program variant is not equivalent to, nor even analogous to, performing benefit-cost analysis. The impact of costs is normally indirect: cost purchases services, which in turn affects outcomes. Therefore, cost and treatment intensity should be analyzed jointly. A finding that cost has an effect over and above service type or intensity means that the marginal cost of that service is higher or lower than average.

In addition, cost bears a different relationship to outcomes depending on the nature of the intervention. Since half or more of OJT jobs are likely to be retained at follow-up, OJT wage, the primary cost factor, will have a strong and automatic relationship with post-program earnings. The question of whether higher or lower reimbursement rates (costs) would improve retention, and thus earnings, for the *same* job is more difficult to address. In the case of classroom training, costs vary more by type of trainer and historical development of each field than by typical wage in the field.

Cost is, of course, a critical issue for managers, and measures of individual treatment costs promise valuable feedback, as long as interpreted carefully.

Program Exit and Job Search

Agency implementation variables discussed earlier measure the availability of various supports at termination. Individual level measures indicate who makes use of which and also provide the basis for aggregate agency level variables.

POSSIBLE MEASURES OF INDIVIDUAL TREATMENT COST

- **For All Interventions:**
  - -- Total direct cost of program activities.
  - -- Total cost of ancillary services received.
- **For Employer-based Interventions:**
  - -- Rate of reimbursement, recorded from the contract.
  - -- Total cost of wage reimbursement.
  - -- Cost of CT associated with employer-based interventions.
- **For Classroom Training Interventions:**
  - -- Cost of tuition and fees.
  - -- Level of needs-based payments per week.
  - -- Total cost of needs-based payments during enrollment.

POSSIBLE MEASURES OF THE PROGRAM EXIT PROCESS

Job Search Services

- Formal enrollment in JSA job club or less extensive job search courses.
- Receipt of less formal job search assistance from JTPA staff.
- Receipt of specific job referrals from trainers or JTPA staff.
- **Participant's job search orientation** (preferably measured at termination).
  - -- Expressed importance of finding or retaining work using skills learned during the JTPA treatment.
  - -- Is job search specific to the program skill area, or broad?
  - -- Is job search geographically local?

Outcomes at Termination

- Standard termination measures already collected should be included for potential use in the analysis of post-termination outcomes.
- If the participant was employed at termination, when did the job begin, and how much was earned by termination? This information allows the follow-up survey to measure only the period after termination, without losing valuable information.
- How much time separates the end of training activities from official termination, and what was the participant's formal program status during that time?

In the case of job search assistance services, it is also useful to measure when the services occurred. If a job search workshop occurred during or prior to training, all participants have its benefit by the time they have to look for work. If the workshop occurs after the end of training, the most successful participants will not enroll because they will have found jobs already.

The expressed importance of working in the training area can be used as a control variable but can also be analyzed as an intermediate program outcome. Those intending such analysis may wish to measure the same variable at enrollment, to allow an estimate of change during training.

Outcomes at termination are important in and of themselves. In addition, they are necessary to complete an accurate picture of the period between the end of program activity and official termination. Analysis is therefore improved by their inclusion during exit interviews.

PART IV.

MEASURING CONTROL VARIABLES

Many of the most important protections against bias are also of interest as implementation and treatment variants, and have been discussed above. Analysis of these measures serves the dual purpose of testing for possible impact on program outcomes and of testing whether their exclusion from multivariate equations biases estimates of other program variants. This is true of implementation variants measuring policies regarding selection criteria, intake procedures, and practices followed in arranging program placements.

The same is true of individual treatment measures such as pre-program competencies or employment barriers and the route into the JTPA training activity. These are of interest in themselves, but they operate primarily as control variables to improve the usefulness to program managers of findings which estimate the impact of program variants.

Other control variables fall into two categories: *individual background characteristics* and the *labor market environment*. These measures are not analyzed in the hopes of improving programs by changing them. Most of them cannot be changed by program managers, or will not be changed since they are part of the program mandate. Their importance is that they are likely:

1. To affect program outcomes,

2. To differ across service providers, program activities, or other program variants, and therefore,

3. To produce biased estimates of program variants of interest to program managers unless they are measured and included in equations estimating those program variants.

PARTICIPANT CHARACTERISTICS

A Cautionary Reminder

Because we can calculate average differences in outcomes between individuals with different background characteristics such as gender or ethnicity, it is tempting to treat such findings as valid estimates of differential impact. They are not. Gross impact data cannot fully address the question: "Does JTPA benefit one group (e.g., men) more (less) than another (e.g., women)?" An equation can show a reliable difference between post-program outcomes for men and women, after control variables are tested. However, we cannot determine whether that difference is created by the program, by labor market discrimination, or by correlated but unmeasured individual differences. Only a net impact design including untreated individuals can generate unbiased estimates of program effects on any one group or on several different groups of participants.

Suggested Measures

The individual background characteristics most likely to affect program outcomes include:

- Inherited characteristics such as gender and ethnicity.
- Previously achieved characteristics such as education level and work experience.
- Life cycle situation, such as marital status and number of dependents.

Some mix of these measures is normally available in Management Information Systems. Where factors known to affect labor market experiences are omitted from MIS files, or where measurement is truncated to distinguish only program eligibles from non-eligibles, MIS files must be augmented. Borus (1979) provides a detailed enumeration of individual background measures found to influence employment status. Here, a few examples of areas where MIS data are often lacking are offered:

- Pre-program employment and earnings records are seldom adequate for use in program analysis.
- For purposes of analysis, education measures should indicate highest degree completed rather than number of years in school, and should not collapse high school graduates together with GED recipients.
- The single parent status, omitted from many JTPA MIS files, is a critical indicator of need and of employment obstacles, making it a key measure of selection bias and of potential labor market outcomes.
- Displaced homemakers should be identified separately, even if the SDA does not target them, because their relative lack of labor market experience has a different meaning than for others.
- Individuals who are young or have for other reasons been in school recently should also be earmarked to clarify the reasons for their low pre-program earnings and employment history.
- MIS files include prior criminal record. However, that measure is inclusive enough to lack power. The variable likely to matter a great deal more is whether an OJT or WEX placement is on work release. Previous analysis found no difference between other ex-offenders and non-offenders, but found work release participants much more difficult to track, and substantially less often retained by their OJT employers (Simpson, 1984a).
- Individuals referred by other agencies may represent cases of multiple employment barriers, making this a potentially valuable measure.

Pre-Program Measures Paralleling Post-Program Measures

Pre-program measures of employment, earnings, and welfare dependency should be as precisely parallel to post-program measures as possible. They may be used in two ways during gross impact analysis. First, these measures may be seen as control variables like any others, indicating the participant's likely success level without any program intervention. Second, they may be used to calculate *change* from the pre-program period to the post-program period.

The special status of these measures comes from their ability to indicate change. For *descriptive* analysis, calculating change is a valuable approach. However, for *differential impact analysis,* the preferable analysis method is not to use change as the dependent variable, but rather to use the post-program outcome as the dependent variables and to enter pre-program status on the same variable as a control variable in the equation.

One difficulty with pre-program measures is establishing their proper time frame. The problem of the "pre-program dip" in earnings and employment has been grappled with in much detail (Bloom, 1982), making clear that information running back as far as three years before the program can be useful. That period is too costly for the design presented in this guide, but a three or six month period before program enrollment is likely to underestimate the long term earning potential of many participants and will fail to distinguish those with temporary problems from others. Each state or SDA must assess its measurement costs against the precision it demands in its program impact analysis. *The pre-program measurement period suggested by this guide is one year prior to application.* Measurement instruments in Appendix E reflect that recommendation.

Some outcome measures such as "employed versus not employed" and "training-related versus other employment," have no analogue in pre-program measures because no single point in time prior to program entry carries the same meaning as the point of follow-up. Pre-program measures are therefore expressed in terms of averages or proportions of the year before application to JTPA. These would be compared to (and subtracted from) the averages or proportions of the three month post-program follow-up period.

Actual pre-program measures will often involve characterizing several brief jobs and then calculating measures expressed in the terms which make them comparable to post-program measures. An example of specific measures using that approach is shown in Appendix E.

CORE PRE-PROGRAM MEASURES COVERING THE PERIOD ONE YEAR BEFORE JTPA

- **Measures of Employment:**
  - -- Percent of time worked.
  - -- Average number of hours worked per week, while working.
- **Measures of Earnings:**
  - -- Average hourly wage rate while working.
  - -- Average total monthly earnings.
- **Measures of Welfare Dependency:**
  - -- Did the participant or dependents receive any aid?
  - -- Percent of time receiving full grant, or partial grant.
  - -- Average total monthly payment.
  - -- Whether receiving payments at the time of application to JTPA.

OPTIONAL PRE-PROGRAM MEASURES:

- Employment measures covering the pre-program year:
  - -- Longest single job held.
  - -- Number of different jobs held.
  - -- Whether any work was in the JTPA training area.
  - -- Did the participant hold any potentially primary labor market job (full-time, non-temporary, with benefits).
- Longest full-time job during the five years before JTPA.
- Reasons for pre-program unemployment:
  - -- Proportion of unemployed time the participant looked for work.
  - -- Laid off from work in a declining industry.
  - -- Fired from one or more jobs.
  - -- Personal reasons (residential move, family change, health).

LABOR MARKET ENVIRONMENT

Measures of labor market characteristics are quite powerful in national studies. They appear to have less effect on training outcomes in a single state or locale. Even so, any study comparing service delivery in more than one geographic area must test the possibility that differences in unemployment levels, average salary levels, or demand for particular types of jobs may affect estimates of differential impact analysis. In addition, when comparisons are made across job titles and across industrial sectors, different placement rates are certain to reflect variations in labor demand and supply across fields and industries.

Aside from census and employment data, the availability of labor market measures depends primarily on the role each state has played in developing reliable occupational outlook data. Most SDAs have compiled this type of information during their planning periods.

IDENTIFICATION OF MEMBERSHIP CATEGORIES

The simplest, but most important, form of data any differential impact analysis requires is a *set of code numbers which identify membership in various categories.* Each unit to be identified during analysis must have a unique identification number. These are required for three reasons:

- They are necessary to the construction of either/or membership variables.
- Some of these identifiers are required in order to merge data from different sources, allowing the construction of data sets that include the full range of test and control variables, and allowing inexpensive once-only measures to be integrated into individual level analysis.
- Some identifiers are necessary to organize the data set and to know what original records to consult in cases where errors on the computer file must be corrected.

At least the identifiers listed below should be included in any analysis. The precise nature of each identifier depends on the common practice in the state or SDAs mounting the analysis effort.

Participant Identifiers

Participant identifiers are the basic data file organizing unit and are also necessary in order to merge data from MIS files, follow-up interviews, and individual treatment records.

The best participant identifier is social security number. These are unique and they are normally required if official data such as UI, welfare, or criminal justice data, are to be combined with follow-up data. If SDA records are organized via any other system of identifiers, these can be employed as long as they are unique across all participants from all agencies to be sampled.

Employer Identifiers

Employer data must be collected under an identifying code which is also recorded on the participant's file. In this way, the appropriate employer follow-up data and/or once-only employer measures may be added to the files of each individual.

- If employers have JTPA identifiers sequenced within year or within agency, codes which are unique across the entire sample must be developed. This can be done by combining identifiers between and within units.

- If agencies have not yet developed employer identification codes, they will find them extremely useful for organizing employer relations and marketing, and for assessing use patterns and retention track records of participating employers.

Classroom Trainer Identifiers

If special data are collected on classroom trainers, they must be catalogued under identifiers also included on participant records to allow data merging. In addition, trainers enrolling a sufficient number of participants in a sample may be tested using either/or membership variables, if each trainer has a unique identifier.

Training Field

The field in which participants trained or gained work experience should be identified, allowing:

1. The description of outcomes by field,

2. The construction of either/or membership variables where the number of cases allows, and, where desired,

3. The introduction of labor market data tied to occupation or training field.

SDA Identifiers

When an analysis combines SDAs, each must be uniquely identified in order to test for differences in outcomes produced by each and in order to add labor market data to individual computer files.

Subcontractor Identifiers

Where subcontractors are used, they must be identified just as SDAs must be. If subcontractors are numbered within each SDA, unique identifiers can be formed by combining SDA and Subcontractor I.D. numbers.

Program Officer Identifier

In systems which assign primary responsibility for each participant to a particular staff member, that staff member may be identified on the participant's data file. This allows a more specific analysis of treatment variants, but should be avoided if there is any fear among staff members that they are being placed in jeopardy.

RECORDING DATE INFORMATION

The simplest reliable way to calculate time periods such as lag between eligibility and enrollment is to record the date of each event: eligibility, enrollment, treatment start, planned treatment end, actual treatment end, termination, and follow-up. If dates are expressed in the correct units, time periods can then be calculated by subtracting one from the other.

The units into which dates and time are coded depend on desired measurement specificity. Measurement expressed in months is too blunt for some treatments, but measures expressed in days are unwieldy for work history questions. A reasonable compromise is to produce a reference sheet showing the sequenced number of each calendar week from 1 to 52, and to use these week numbers to record all dates. This makes all measures of dates compatible and makes calculation of time periods easy, requiring only an adjustment for periods which span calendar years. (See measures in Appendix E.)

CHAPTER 9
EMPLOYER MEASURES

Chapter 9. Employer Measures

Job training programs have impact on employers as well as participants. With the advent of JTPA and the expansion of private sector involvement, interest in measuring employer benefits has risen. The popularity of this issue among service providers is no doubt connected to a concern for marketing JTPA services and products to employers. The perceptions of OJT or WEX employers are useful to indicate which program approaches are relatively effective, which are distasteful to employers, and what steps might encourage or discourage future participation by employers (e.g., Minnesota, 1979; Simpson, 1984b).

In addition, employers may be viewed as direct beneficiaries of the job training system and, in some cases, as incurring costs of providing services to that system. This viewpoint directs the design and measurement suggested in this guide. That is, measures focus on employer costs and benefits only, without assigning portions of those costs or benefits to others, even where they are in fact shared or passed on by employers.

SEPARATING EMPLOYER OUTCOMES FROM OUTCOMES FOR OTHERS

Employers occupy two roles pivotal to the success of the JTPA program -- they employ former JTPA participants and, in some cases, they provide services to JTPA participants. It is therefore difficult to separate benefits to employers from benefits to participants and to society. When a placement works out well, all benefit. When an employer provides training, the participant can become more employable (either within the firm, or generally) and the employer can gain a more productive worker. Similarly, the wage subsidy employers receive is rewarding to the employer and also to the participant, who receives full pay for a period of partially subsidized work.

However, thinking about measures of employer costs and benefits requires a shift in focus to the costs and benefits to employers alone, independent of their implications for participants, taxpayers, or the program itself. These others are important, but separate issues. This guide recommends measuring employer outcomes independently from other outcomes and developing interpretations of the relationships among employer outcomes and other outcomes later, during and after analysis. The measures suggested here reflect that decision.

PART I.

THE NATURE OF EMPLOYER OUTCOMES

Employer outcomes are not specified in the JTPA legislation. Nor is there a long tradition of past research focusing on them and defining them. One must therefore begin by exploring the possibilities. We begin by discussing two major limits which direct measurement design in this case: the lack of prior development in this area of research, and the inability of gross impact research to estimate net impacts.

MEASUREMENT STRATEGY IN A NEW AREA OF STUDY

As a relatively new area of study, employer benefits cannot be measured definitively. It is possible to specify a range of probable benefits and costs, but too little is known about each or about their relative worth to employers to develop a precise accounting of costs and benefits which is meaningful. Some of these costs and benefits, such as the OJT wage subsidy, can be expressed in precise monetary terms. However, others may be equally important to analyze but impossible to quantify or even to conceptualize clearly. For example:

- The major sign that hiring a JTPA participant was rewarding to an employer is an employer's decision to retain the participant. However, it would be complex to estimate how far above or below a breakeven point (a point of indifference, neither costly nor rewarding) each placement falls. Without that information, we cannot claim that retention of more than 50% of participants represents an average net benefit.

- The employer's perception that a JTPA participant may have serious problems not easily observed before hire, is impossible to quantify in precise monetary terms but is a very important cost for many employers (Simpson, 1984b).[1]

- The provision of training is costly to employers. Assigning quantitative values to employer training is essentially arbitrary, even when done by service providers. It is difficult to assess how much training occurs. Further, we do not know with any precision what proportion of training would be offered to all new employees, regardless of JTPA involvement or what proportion of that training is so specific to the particular employer that it binds the worker to that job rather than transferring to other employment situations. The former type is much less costly to employers than the latter.

- Other elements of the JTPA program are even more difficult to measure precisely. It is even unclear whether they act as costs or benefits. For example:

    -- Employee screening can be a service to employers. During CETA OJT, some employers saw agency screening as beneficial, but others saw it as costly, based on their experience with the participants referred to them (Simpson, 1984b).

    -- Working with the disadvantaged is typically assumed to be one cost to employers. Yet one study of CETA OJT employers found over one-tenth listing the knowledge that "you are helping others with need" as *the major reason* for participating in OJT (Simpson, 1984b).

- Similarly, performance on the job by JTPA participants or former participants can represent a major cost or benefit to an employer. Yet, in any specific case, we must measure whether, as well as to what degree, worker characteristics are costly or beneficial. Further, we know relatively little about which worker characteristics are most important to measure -- that is, which are most costly or rewarding to employers.

We face these measurement challenges primarily because this area has been analyzed relatively little by past work, and because many of the most important costs and rewards to employers are inherently perceptual and therefore not susceptible to monetary quantification. One measurement option in this situation is to focus on the few measures which can be quantified, making whatever assumptions about the nature of these costs and benefits are necessary to such quantification. (For an example of this approach tailored to JTPA, see Stromsdorfer, 1986.)

---

[1] Quantitative estimates could be derived from large scale research designed for that expressed purpose. No such estimates now exist for us to build on, and the gross impact approach suggested here is not suited to the development of such estimates.

That option is not appropriate to the gross impact approach. This guide adopts the second option: specifying a range of possible costs and benefits to be measured in whatever form is most appropriate to each. This often means measuring employers' perceptions, using survey answer scales which can be analyzed quantitatively but cannot be transformed into monetary terms. The emphasis in this approach is on identifying multiple measures and investigating the extent to which each is perceived by employers to act as a cost or a benefit, as well as estimating the importance of each. This approach of emphasizing employer perceptions is unusual for studies of costs and benefits. Indeed, *this guide does not recommend attempting a formal benefit-cost analysis of employer benefits.* Instead, much is to be gained by *exploring* the nature and meaning of costs and benefits to employers. For example:

- We learn what aspects of JTPA are most costly and most rewarding to employers.

- The knowledge we gain is expressed in terms which not only help us to build a better understanding of employer outcomes, but also speak to JTPA agencies' need to market JTPA to employers -- that is, in terms of employers' perceptions.

- We can analyze whether particular types of employers have different perceptions of the costs and benefits of JTPA, and whether these different perceptions are associated with greater or lesser program success for participants.

- We can test our ideas about the ways in which JTPA is rewarding or costly to employers. Rather than assuming that particular JTPA services such as client screening are costly or rewarding to employers, we can examine the extent to which the implementation of these services increases or decreases the rewards or costs perceived by employers.

- We can examine the relationships among various costs and benefits. Do they represent trade-offs, with some employers or service providers producing one mix of benefits and costs, while others produce a quite different mix? Or are rewards of various types mutually reinforcing, so that some settings are more rewarding than others on many ?

ESTIMATING COSTS AND BENEFITS USING THE GROSS IMPACT APPROACH

Within the limits of the gross impact approach, true net impact of JTPA on employers cannot be estimated. That would require comparisons with employers hiring non-JTPA participants. Nevertheless, we can say something about level of cost or benefit of participating in JTPA or of hiring a former JTPA participant.

To accomplish this, we can use measures which *ask employers to give us their estimates of their costs and benefits* of participation. The measures suggested in this guide accomplish this by specifying a *breakeven point* for each measure of cost or benefit, and asking employers to report whether their experience with JTPA or with specific JTPA participants placed with them fell above or below that point. The breakeven point differs by type of measure and is discussed below. The strategy in each case is to express the measure in terms simulating true or perceived net cost or benefit, by wording the measure in terms of breakeven point, and offering responses on either side of that point. Precise degree of cost versus benefit will not be estimated. Instead, the measures suggested here will emphasize clearly establishing the direction (cost or benefit) from breakeven for each measure, and more loosely estimating the distance from breakeven. (See Appendix E.)

TRADE-OFFS AMONG DIFFERENT COSTS AND BENEFITS

It is the goal of true benefit-cost analysis to estimate a single quantity showing the overall net cost or benefit of some phenomenon such as hiring a former JTPA participant. This guide does not speak to this possibility except to note that *this gross impact design is not a benefit-cost* *approach*, and that in the case of employer outcomes, such a design may be impossible to implement at this point in the development of the field. Nevertheless, measuring various costs and benefits to employers raises the question of how to combine them -- of how we can assess whether the benefits we identify offset the costs we identify.

Similarly, we face the question of how to evaluate employer benefits in conjunction with costs and benefits to others. Some benefits to employers are passed along to participants; others are not. For example, training on the job may or may not be general enough to transfer to other employment contexts. Some benefits to employers, participant wage subsidies and JTPA staff screening time, also represent costs to the taxpayer, although these may be offset by program successes resulting from employer placements. Similarly, employer risk-taking and training activity may represent costs to the employer and benefits to the participant and to society.

To assess with precision how much the costs and benefits to employers offset each other or are offset by being passed along to others would be extremely difficult. To do so within the constraints of gross impact design and given the lack of prior development in this area of research is impossible. We possess little information identifying what the costs and the benefits to employers are, and even less about the nature of tradeoffs among them. Nevertheless, some information can be gained during analysis.

First, it is possible to *analyze the association among different measures of cost and benefit.* Are employer costs of providing services higher where benefits, such as the subsidy to participants' wages, are higher? Do employers who receive high levels of one type of benefit tend to receive less of others, or is JTPA implementation such that some agencies reward employers more than others across the board? The answers to these questions can offer some guidance to interpretations of program service levels and the value of expenditures on employer-based program activities compared to others.

Second, *employers can be asked to indicate how important they perceive different costs and benefits to be.* While this approach falls short of precision, it allows analysts to prioritize employer outcomes. It is clearly more important to overall employer costs and benefits that outcomes be perceived as positive on major outcomes than that they be positive on irrelevant outcomes.

Third, we can analyze the degree of association between each measure of cost or benefit and an overall criterion of net success which appears valid on its face. For example, participating employers can be asked whether they intend to participate again. Termination employers can be asked whether they have retained the former participant whom they hired, and whether they intend to retain that individual in the future. The degree of association between each measure of cost or benefit and the overall criterion can be interpreted as a proxy indicator of the relative power of that cost or benefit to determine net benefit-cost.[2] Such an approach is also useful to analysis which has the goal of improving employer marketing or improving participant success levels. Employer marketing is improved by knowing what aspects of the program must be improved in order to retain the interest of participating employers. Participant outcomes can be improved by learning what aspects of the service provider's relations with the employer have the greatest impact on eventual program success for participants (i.e., retaining employment).

SPECIFYING OUTCOMES FOR TERMINATION EMPLOYERS

Employers whose association with JTPA begins only when a former participant is hired experience certain costs or benefits which are associated with JTPA. These are the qualities of the former participant as an employee. The question for these "termination employers" is whether the new JTPA-trained employee will function as well in the job as other appropriately trained new workers. *There is no reason to expect JTPA participants to be better trained than others.* The goal for participants is *elimination of their previous deficit.* For employers, the goal is to *hire the most satisfactory worker* available, regardless of the involvement of JTPA.

---

[2] A proxy is a substitute indicator of some phenomenon. It is associated with the phenomenon but is not a true measure of the phenomenon itself.

During periods of less than full employment, availability of appropriate job applicants is rarely an issue for employers who hire JTPA participants. They choose one of numerous applicants who appear to have adequate personal and work qualifications for the job in question. Variation above or below *average new hire* indicates the primary benefit or cost to employers of hiring a JTPA participant. For example, participants whose weak job history has been fully counterbalanced by JTPA training may be superior workers, representing a JTPA benefit to the employer. However, participants whose assistance from JTPA was unable to reduce continuing personal problems may be less adequate workers than they appeared at hire, representing a cost to the employer.

In each of these cases, it is difficult to completely separate the costs and the benefits employers might experience. Each area of cost or benefit can be considered to represent one dimension, or scale, such as how many days' work will the new employee miss during the first month, or how much employer training will be required before the worker becomes productive. Cost and benefit represent two ends of each scale. The breakeven point lies at the point on the scale which represents the average new hire for that job in that labor market, as perceived by the employer. If the average new hire misses four days' work per month, hiring a former JTPA participant who misses an average of two days represents a benefit to the employer.

The measures included in this guide's ready-to-use surveys employ a simple five point scale to measure this type of employer benefits.[3] (See Appendix E.) For each of several characteristics, the employer is asked to indicate whether, compared "...with the average workers you have hired or could hire for this same job..." the former participant they hired is:

much better,
a little better,
about the same,
a little worse, or
much worse.

In the case of any one participant, job performance may be better or worse than average for reasons unrelated to JTPA participation or referral. However, if, over a large number of employer interviews, the average JTPA hire proves to be more satisfactory to employers than their average non-JTPA hires, we have reason to claim a role for JTPA in producing that benefit to employers.

OUTCOMES FOR PARTICIPATING EMPLOYERS

For employers who participated in the JTPA program by providing on the job placements or work experience, the most obvious benefit is the subsidy to participant wages. By far the most common reason employers give for participating in OJT is the subsidy. Other reasons commonly reported include eliminating the need to screen large numbers of applicants, the ability to expand or to stabilize without mounting the full cost for the new employee, and satisfaction at being able to assist deserving individuals (Simpson, 1984b). Commonly reported costs include the time and supervision required to train, the potential of greater than average work time lost to personal or family problems, the possibility that maximum performance after training will not match that of other employees, and the possibility that JTPA employees might turn over faster than others would.

---

[3] Written surveys may use finer gradations, but for telephone surveys, more than five response options becomes awkward.

Some of these issues represent only degrees of cost or degrees of benefit. The breakeven point for the OJT wage subsidy is zero; it cannot be costly in and of itself. (The costs associated with accepting it should be measured separately). Similarly, paperwork requirements cannot be seen as benefits; they are always costly to some degree.

Other benefits and costs to participating employers are meaningful only when a breakeven point is defined in comparison to typical employees who would be hired were it not for the JTPA program. The two major costs JTPA wishes participating employers to accept are hiring individuals who appear to be less qualified for the job than typical non-JTPA hires, and providing extra training beyond that required by typical non-JTPA hires. The issue is not, for example, whether the OJT employer loses five or ten weeks of productive time during training, but whether the *difference in training time* for typical non-OJT hires versus the OJT hire is zero, five, or ten weeks. A difference of zero weeks represents a breakeven point on that particular measure.

The difference between JTPA and other hires is measured in the same ways for participating employers as for termination employers. The difference arises only during interpretation. The assumption is that participating employers receive a wage subsidy as an incentive for hiring or training a participant who is less qualified than typical non-JTPA hires. The measurement strategy suggested here is that measures establish a breakeven point at zero difference between JTPA and non-JTPA hires, just as for termination employers. Analysts may then determine whether they are satisfied with the differences employers report between JTPA participants and other hires, given the JTPA reimbursement they receive.

Studies of CETA OJT employers have found OJT hires perceived to be more similar to other hires than might have been anticipated (Simpson, 1984b; Minnesota, 1979). Apparently, all those who apply for the secondary labor market positions to which some OJT participants are assigned are in some degree "disadvantaged." This type of finding illustrates the importance of simulating net comparisons between JTPA and other hires.

CONSTRAINT, INVESTMENT, AND WINDFALL AMONG PARTICIPATING EMPLOYERS

It is not always the case that employer-based programs work best for participants when employers receive maximum benefit from participation. The issues discussed below focus on this issue.

One element of employer costs is the degree of constraint experienced by the employer. O'Neil's (1982) analysis of employer hesitance to use Targeted Jobs Tax Credits demonstrates that the sheer fact of being constrained can be costly to employers. One way in which employers' perceived cost/benefit ratio from OJT participation have been improved since early CETA is through the reduction of costs due to constraints from paperwork and insuring the ability to reject referrals. Earlier programs had protected their right to serve participants with greatest need, but in so doing had raised employer costs above the threshold allowing participation.

At the other extreme, far from being overly constrained, some participating employers use the federal wage subsidy without incurring the cost of providing any services. In-depth interviews with CETA OJT employers located some who explicitly stated that they provided no training and refused to alter their hiring practices at all, choosing instead simply to gather the extra income wherever one of their new hires happened to be OJT eligible (Simpson, 1984b).

This is a problem regarding the level of service provided by JTPA, rather than an issue of employer benefits as such. Payment for non-service is certainly beneficial to the employer involved, but it represents an unintended windfall, not an intended employer benefit. This makes the issue doubly important to include in any analysis of employer outcomes.

An issue intermediate to the above problems of constraint and windfall is identified by economic theory regarding non-subsidized on-the-job training (Maranto and Rodgers, 1984; Hoffman, 1982). Employers always engage in *introductory OJT*, specific to the firm and to the job. This training represents part of the employer's investment in hiring any new employee, and part of an employee's investment in developing continued employment. In addition, more general training is typically made available for selected employees. Since the *training represents an investment on the part of the employer*, it is normally reserved for the most promising employees, who will be retained and promoted, and who will repay the investment through higher productivity.

The typical sequence is, therefore: hire, decide to retain, decide to invest in training, train, and retain as planned. The subsidized OJT situation differs from this typical sequence in two ways: the *training occurs before the decision to retain,* and the *training may not be the result of a decision to invest* in training. The first of these departures represents the core nature of the OJT program. The subsidy purchases an opportunity to work and train for a period prior to the actual hiring decision. The second issue is more complex.

If the total cost of training a JTPA participant is greater than the income derived from the wage subsidy, the employer must decide to invest in training, which in turn implies a commitment to hire if possible, so as not to waste the investment. If, however, for some reason such as the need to enlarge the OJT program, an SDA offers subsidies equal to or larger than the employer's cost, the employer may decide to participate without ever having deciding to invest in the participant. The reason may be kind ("Now I can afford to help this person") or hard nosed ("I make more money hiring OJT even if I increase turnover by letting them go after the contract ends"). The result in either case is that the employer experiences less pressure to make certain the training is good and the hire permanent. For the program to work well, the employer's perceived cost/benefit ratio must be good enough to encourage participation in the program but not so good as to reverse the economic incentive of the OJT investment.

SOURCES OF EMPLOYER COST AND BENEFIT MEASURES

Most employer benefits studied using the gross impact approach are measurable only through interviewing employers. (See Chapter 5.) One *prima facia* benefit becomes available with follow-up data from any source: whether the JTPA participant proved to be desirable enough and reliable enough to remain employed at follow-up. In the case of participating employers, some direct financial benefits are available in agency records, although typically not as part of a computerized MIS. These include the wage subsidy, and support for other training useful to the employer. Support services are presumably a subsidy to the participant only, and not a benefit to the employer. Other outcomes such as "value added" to the firm are not considered in this guide but are measured by earnings.

PART II.

MEASURES OF EMPLOYER COSTS AND BENEFITS

Use of the terms *cost* and *benefit* require that one point made earlier be repeated. The measures suggested below do not cumulate to a precise benefit-cost figure. This is not benefit-cost analysis. It is systematic description of employers' perceptions of their costs and benefits. It identifies areas of potential cost or benefit and checks those against the reports of employers. Analysts may wish to try out combinations of measures, but should be aware of the problem of double counting. No effort should be made to express the employer costs or benefits developed here in quantified financial terms because each is not unique from the other.

OUTCOMES WHICH APPLY TO TERMINATION AND PARTICIPATING EMPLOYERS

For all employers, including those who become involved with JTPA only by hiring a former participant, the job behavior of the participant represents one basic set of costs and/or benefits. Assuming that the JTPA intervention was successful in moving an individual up the employability queue enough to produce this hire, how successful was the intervention at providing this employer with a *productive and tractable employee?* This is the key benefit to the firm and to the society as a whole.

For each of these measures, the employer's perceived breakeven point is the average of the typical non-JTPA workers who could have been hired into the same job under the same economic conditions, as discussed above. Several measures of this type which have proved useful in previous analysis are shown on the following page and in Appendix E. These include indicators of skill and productivity, ease of supervision, and job adjustment. In each case, ratings higher than the average non-JTPA hire indicates employer benefits.

A second group of outcomes recognizes that the single most important benefit for termination employers is the ability to retain the worker they hired. *Retention* implies that the worker is productive and adjusted, and also wishes to remain employed. Unless non-retention resulted from cutbacks forced by declining business, laying off a trained, productive worker indicates a cost to the employer. Whether these costs occur because participants perform poorly or because they quit is also valuable to explore. When workers are not productive, of course, failure to retain the participant represents a reduction of costs.

MEASURES OF PERFORMANCE ON THE JOB, COMPARED TO NON-JTPA WORKERS

- Skill level:
  - -- How fully trained when hired.
  - -- How rapidly required job skills were learned.
- Job performance:
  - -- Overall productivity.
  - -- Getting work done quickly.
  - -- Working well independently.
- Supervision ease and work habits:
  - -- Overall ease of supervision.
  - -- Following directions well.
  - -- Being willing to do extra work.
  - -- Being enthusiastic on the job.
  - -- Being honest and reliable.
- Personal adjustment on the job:
  - -- Getting along well with others.
  - -- Being able to handle job stresses.
  - -- Keeping personal life from interferring with work.

JOB RETENTION AND REASONS FOR NON-RETENTION

- Is the participant still employed?
- If not retained:
  - -- How long did the employment last?
  - -- Did the job end because of cutbacks (a cost to the participant, not the employer), did the participant quit, or was the participant fired?
  - -- Does the reason for the quit or firing indicate inadequate job performance, poor work habits, or personal problems?
- If retained:
  - -- How likely is retention for the next year (or other period?)
  - -- Is a promotion or raise, likely within the next year?

OUTCOMES FOR PARTICIPATING EMPLOYERS ONLY

Outcomes Measured Through Agency Records

The most basic benefits which accrue to participating employers are financial, and may be recorded directly from JTPA contracts. This form of measurement is preferable because it is highly reliable, it indicates both planned and actual expenditures, and it avoids the awkwardness of asking about money during telephone interviews. In addition, agencies may be able to estimate the amount of screening and referral time they provided, thereby offsetting employer hiring costs. The question of how effective the screening was is separate, and must be measured during employer interviews.

EMPLOYER BENEFITS WHICH CAN BE RECORDED FROM AGENCY RECORDS

- Total amount of wage reimbursement to the employer, planned and actual.
- Proportion of wages reimbursed.
- Time period of reimbursement.
- Total amount spent to provide classroom training supportive to the employer placement, if requested by employers.
- Amount of screening and referral time provided by the agency, offsetting employer hiring costs.
- Number of JTPA participants placed with this employer per year.

Outcomes Measured through Follow-up Surveys

In addition to measures listed earlier for all employers, participating employers incur training and other costs during their contracts, and also experience the potential costs and benefits of working with JTPA agencies. These may be measured through employer follow-ups, in the form of employer reports of their activities or employer perceptions of JTPA.

The Cost of Training Investments. One category of employer cost includes direct investments in training activities. In this case, employer measures can serve two purposes:

- They can indicate employer cost or benefit by estimating the amount of training effort expended for the JTPA participant compared to the average non-JTPA hire.
- They can indicate how intense the training intervention was for each participant, as part of a differential impact analysis of participant outcomes.

Although only the first of these is germane to this chapter, the measures suggested below are useful for both purposes. It is primarily for this reason that both absolute measures and measures relative to non-JTPA workers are included. In addition, employers can estimate the proportion of their training which is specific to that one job in that particular firm. This estimate can help us address one case in which the long range costs of training to an employer may be greater or smaller depending on whether the training ties an employee to the firm or makes the employee more employable with all firms.

MEASURES OF EMPLOYER INVESTMENT DURING CONTRACTS WITH JTPA.

Measures Comparing JTPA to Average Non-JTPA Hires:

- Length of training provided as a proportion of typical non-JTPA training time.
- Employer's estimate of difficulty or intensity of training compared to training non-JTPA employees.
- Whether special training methods have been developed for JTPA hires.

Measures of the Absolute Amount of Training:

- Length of training period.
- Average number of hours per week devoted to training during the training period.
- Use of formal training methods (assigned reading material, in-house training kits, classes) in addition to purely informal learning.
- Proportion of training which applies to this employer and job only, versus training which is transferable to other positions.

The Cost of Accepting Risk or Constraint. Although particular JTPA participants may prove to be ideal workers, it is in the nature of a program which offers subsidies in exchange for hiring particular individuals that some risk is implied. The participant *could* be a poor worker, an alcoholic, or a thief. This risk may loom larger than the costs experienced because particular participants perform poorly. The JTPA agency could also attempt to constrain the employer's behavior, or unexpected paperwork demands could develop. These *possibilities* may be costly in employers' perceptions.

At the other extreme, the employer could reduce risk by retaining control over the hiring process. In the most extreme case, employers make firm hiring decisions before sending their new employees to apply for the OJT subsidy.

MEASURES OF EMPLOYER COST VIA RISK TAKING

- How much influence over hiring did the employer transfer to JTPA?
    - -- None, hiring was firm before eligibility was determined.
    - -- Little, referred to JTPA before firm hire, or interviewed JTPA eligible along with non-JTPA applicants.
    - -- Much, accepted only JTPA referrals.
- Employer's perceived risk of hiring an especially difficult case.
- Employer's perception of pressure to retain participants or of other constraints in order to participate.
- Degree to which employers perceive that paperwork is constraining.

Employers Direct Evaluations of Costs or Benefits of Participation. Most of the measures suggested thus far in this chapter are indirect, in that they ask employers to rate a particular JTPA participant or placement experience. This approach has the value that it defuses possible employer concerns about being evaluated: it is the participant who is being evaluated. Employer responses can be interpreted evaluatively when they are aggregated, but not individually.

In addition, certain employer outcomes are best estimated in a direct form. Some of these are evaluations of JTPA or JTPA services. Others are specific aspects of participating in JTPA which employers can be asked to rate as costs or benefits. The survey included in Appendix E asks whether each is a major or a minor cost or benefit. Some of the measures listed below, the subsidy to wages and the feeling of helping others, probably cannot represent costs. Others, such as JTPA screening of applicants, may be costs or benefits, depending on their quality. However, placing all in a list of questions asked using the same answer format is efficient and employers will tend to feel more able to express criticism honestly after answering a clearly positive question such as whether the wage reimbursement is beneficial.

DIRECT EMPLOYER REPORTS OF COSTS AND BENEFITS

- Evaluations of the JTPA program and service providing agencies.
    - -- JTPA administrative efficiency and responsiveness.
    - -- The JTPA representative working with the employer.
    - -- JTPA screening of applicants sent to employers (where applicable).
    - -- Qualifications of JTPA job applicants, on the whole.
    - -- The overall JTPA program.
- Intention to participate in JTPA again.
- Perceptions of various aspects of participation as costs or benefits, of major or minor proportions.
    - -- The wage reimbursement.
    - -- JTPA screening of applicants.
    - -- Reimbursement for courses during training, if applicable.
    - -- Enlarging or stabilizing the employer's work force.
    - -- The feeling of helping people with need.
    - -- Program participation, overall.

CONCLUDING NOTE

A variety of measures have been suggested in this chapter (and in Appendix E), in the belief that the employment and training field is in only the first stages of developing firm indicators of employer outcomes. Most of these measures have been utilized in previous research (Simpson, 1984a; 1984b). Nonetheless, they are suggested here primarily as a stimulus to the development of work in this area. Analysis which tests the interrelation of such measures or their importance in explaining which employers agree to future participation will be of immediate benefit to managers conducting the analysis and of long term benefit to the development of this area of study.

SECTION IV.

ANALYSIS

Chapter 10. Analysis Procedures

This chapter makes no attempt to provide instruction in the use of statistics. It begins with a brief overview of approaches to analyzing descriptive analysis, differential impact analysis, and employer outcomes. It then discusses selected issues regarding how to analyze data from the survey instruments included in Appendix E. These include: levels of measurement and the most common statistics used with each, performing multivariate analysis using a dichotomous (dummy) dependent variable, testing dummy variables and interaction terms, use of standardized versus unstandardized regression coefficients, analysis of pre- to post-program change, strategies for moving through analysis of a large number of variables, and merging data from different sources.

One other topic should be mentioned because readers may refer to this chapter for information on it. A convenient first step in setting up relatively large data sets for analysis is listing all variables measured, along with identifying information and rules for transforming the raw data into usable form. This guide includes a table showing that information for the ready-made survey instruments included with the guide. Both the surveys and the table are included as Appendix E. For those who adopt the ready-made surveys, this table reduces tedium considerably. For others, the table may serve as illustration.

OVERVIEW OF GENERAL APPROACHES TO ANALYSIS

Descriptive analysis involves quite basic statistical tools. As discussed in Chapter 3, the value of descriptive analysis rests more on the thought which goes into the questions which the analyst asks than on statistical sophistication. Descriptive analysis begins with univariate (one-variable) averages or percentage distributions. Beyond that, bivariate (two variable) associations can be calculated, as long as the analyst bears in mind that descriptive associations can be produced by many factors other than the two being analyzed. Exhibit 10.1 describes conditions under which different bivariate statistics are most appropriate.

Differential impact analysis can be performed satisfactorily with standard multiple regression techniques, except for particular situations discussed below. The strategy of multivariate analysis is straightforward. One outcome is analyzed, with multiple potential influences tested against each other to determine whether the policy variables of interest affect the outcome after taking into account the possible effects of other factors. However, the statistical techniques required to implement that strategy require specialized training. Part II of this chapter discusses several specialized topics likely to arise during differential impact analysis of JTPA outcomes. However, these discussions assume prior background with multivariate analysis.

Analysis of *Employer costs and benefits* is not directed by a single analysis strategy. Basic descriptive analysis is especially valuable to this relatively new area of study. In particular, employer perceptions have considerable prima facia value to the JTPA programs to which employers are so central. Most measures of employer outcomes which are suggested in this guide are structured so that descriptive percentage responses are meaningful. In addition, the relatively uncharted nature of employer outcomes makes descriptive bivariate associations valuable to identify. It is useful to know how various benefits are related to each other and how costs are related to benefits. In the case of employer outcomes, these associations may now be assumed, but are not confirmed by systematic analysis. Finally, multivariate analysis is also appropriate, for purposes of identifying probable influences on employer outcomes and for testing the influence of employer variables on participant outcomes.

PART I.

ANALYZING VARIOUS TYPES OF MEASURES

LEVELS OF MEASUREMENT

Many statistics are available in various software packages. However, nearly all statistical tests required for descriptive or differential impact analysis can be performed with four basic tools: Chi square, analysis of variance (ANOVA), Pearson correlation, and ordinary least squares (OLS) regression. (Why these are typically adequate is laid out in a highly readable form in Bjornstadt and Knoke, 1980.) Which of these is used depends on the goal of the analysis and on the way in which the variable was measured -- commonly referred to as the level of measurement.

Although statistical analysis should only be conducted by individuals well versed in statistics, this guide offers a set of reminders to help prevent errors of oversight during analysis. In Appendix E, the level of measurement of every variable included in the ready-made surveys is shown. Only those levels of measurement which affect choice of statistical tools are shown. Exhibit 10.1 supplements that listing by suggesting appropriate statistics for different levels of measurement and for different analysis goals.

The critical level of measurement distinction is between ordered and nominal variables. *Ordered* variables are those for which values assigned to each category of the variable form a logically defensible sequence from smaller to larger, lower to higher, etc. Ordered variables include age, level of satisfaction, costs, ratings on various descriptive scales, and the like.

Variables which cannot be ordered are termed *nominal* variables. The categories of variables like marital status or SDAs identification codes cannot be placed in a meaningful hierarchy or sequence. The results of tests that require ordered variables are meaningless for nominal variables such as these.

Dichotomous variables, those taking only two values such as "yes" and "no," occupy a special status in that they are by definition ordered, even when they appear non-orderable, as in the example of which SDA one enrolled through. If only two values are included, the variable can be expressed as a yes/no question. In the case of one SDA versus others, the variable becomes "Did this participant enroll through SDA #1?" The responses, "yes" and "no" are interpretable as ordered, with yes greater than no. It is this quality of dichotomies which makes dummy variables especially useful, as discussed later in this chapter.

Statistical assumptions vary somewhat for dependent (outcome) variables versus independent (predictor, explanatory) variables. Therefore, the choice of statistical tools depends on the level of measurement for each. Exhibit 10.1 reflects this requirement. Analysis goals are separated into bivariate (two variable) and multivariate (one dependent variable, more than one independent variable) cases, and level of measurement is indicated for both independent and dependent variables.

EXHIBIT 10.1 SUGGESTED STATISTICS FOR DIFFERENT LEVELS OF MEASUREMENT

	Type of Variable:		
Analysis Goal	Independent	Dependent	Suggested Statistic
A. Bivariate	Ordered	Ordered	Correlation[1]
B. Bivariate	Ordered	Dichotomous	Treat as type C[2]
C. Bivariate	Dichotomous or Nominal	Ordered	ANOVA (F test)[3]
D. Bivariate	Dichotomous or Nominal	Dichotomous or Nominal	Chi Square
E. Multivariate	Ordered	Ordered	OLS regression
F. Multivariate	Dichotomous	Ordered	OLS regression
G. Multivariate	Ordered	Dichotomous	Varies. See below
H. Multivariate	Any	Nominal	None, or ECTA[4]
I. Multivariate	Nominal	Any	Transform to dummy variables

---

[1] If ordered variables contain few (3-6) categories, it may also be advisable to observe relationships in tabular form. However, the chi square statistic would typically underestimate the likelihood of a reliable relationship because it ignores information on order.

[2] In this case it is convenient to treat the independent variable as the dependent, and vice versa, so that ANOVA may be used. ANOVA allows the ordered variable to take a large number of values and makes full use of the ordered nature of one variable in the pair, arbitrarily named the dependent variable.

[3] In the dichotomous case, the t-test is equivalent to the F test used in ANOVA.

[4] If available to the analyst, recent developments by Goodman (1972) make limited multivariate analysis of nominal variables possible. (See also Davis, 1974.) Goodman's program is named ECTA (Everyman's Contingency Table Analysis). SPSS-x has also installed a version.

PERFORMING MULTIVARIATE ANALYSIS WITH A DICHOTOMOUS DEPENDENT VARIABLE

Every statistic is developed on the basis of mathematical assumptions. In the case of OLS regression, many of the original restrictive assumptions have proven unnecessary. That is, the statistic is highly robust (Bjornstadt and Carter, 1974). Even so, in the case where the dependent variable is dichotomous or is highly skewed (unevenly distributed), assumptions are violated severely and error can result from use of OLS regression.

Happily, recent work with statistics based on log-linear transformations of dependent variables and using maximum likelihood chi square goodness of fit tests rather than variance explained tests of reliability avoid the problems faced by OLS regression. This means that there exist appropriate conservative multivariate methods to analyze dichotomies such as whether or not participants are employed at follow-up. It also means that OLS regression tests have been compared with these more conservative methods. The result is that we can be quite certain when the more conservative, but also less convenient, methods must be used and when the simpler OLS form of analysis is appropriate. (See Knoke, 1975; Goodman, 1976; Gillespi, 1977.) In the context of this guide, the following summary guidelines will serve:

- If a dichotomous dependent variable is split relatively evenly (between 75% / 25% and 25% / 75%) OLS regression may be used.
- If the split is less even than the above, OLS regression might be used to explore, but should be avoided for final tests to be reported.
- If OLS regression cannot be used and all or many independent variables are ordered, Logit or Probit transformations of the dependent variable are advisable.
- If OLS regression cannot be used and many independent variables are dichotomous or nominal, Goodman's Multiway Contingency Table Analysis may be used.

CONSTRUCTING AND TESTING "DUMMY" VARIABLES

Some factors that are extremely valuable to include as independent variables in OLS regression analysis can be measured only as nominal variables. This may be accomplished by transforming the nominal variable into one or more dichotomies and testing them rather than the original variable. These dichotomies are place holding membership variables called "dummy variables." Since dichotomies may be included in OLS regression runs, dummy variables allow an indirect way of testing the impact of nominal variables. The construction of dummy variables is illustrated in Exhibit 10.2, using the example of marital status. For example, the new variable "Married" takes the value 1 (yes) where the original variable had the value 2 (now married). In all other cases, this new dummy variable takes the value 0 (no).

Once variables are redefined in this way they may be used as independent variables, and, under the conditions discussed above, as dependent variables in OLS regression analyses. This type of variable is inevitably crucial to differential impact analyses because dummy variables provide a mechanism to include specific SDAs or service providers and specific program activities into regression equations.

Regression slopes which result from tests of dummy variables, if statistically reliable, indicate that members of the named group (e.g., OJT participants) are some estimated average amount higher or lower on the outcome in question than the categories not included in the equation. All but one of the dummy variables created from an original nominal variable may be tested in a single equation. If forward stepwise procedures are being used, all may be included.

EXHIBIT 10.2. ORIGINAL AND DUMMY VARIABLE FORMS OF MARITAL STATUS

Original Categories		Four Dummy Variables and Their Values			
		Nev. Married?	Married?	Div./Sep?	Widow?
Never Married	= 1	1 (yes)	0 (no)	0 (no)	0 (no)
Now Married	= 2	0 (no)	1 (yes)	0 (no)	0 (no)
Divorced/Separated	= 3	0 (no)	0 (no)	1 (yes)	0 (no)
Widowed	= 4	0 (no)	0 (no)	0 (no)	1 (yes)

CONSTRUCTING AND TESTING INTERACTION TERMS

One useful type of question for JTPA program managers may be addressed through use of interaction terms: whether particular types of participants experience greater success from some program variants than from others. For example, does CT do better than other treatments at erasing the deficit produced by previous low educational attainment? This kind of hypothesis specifies an interaction. That is, two variables combine to produce a different effect than either alone or both acting independently would produce. To test such an hypothesis, one must construct an interaction term by combining two original variables. That new term is then tested along with the two original variables in a single equation.

Interaction terms can become difficult to interpret. However, one form is nearly always manageable. Its construction begins with a dummy variable scored 1 and 0, and multiplies that variable times a second variable of interest. Carrying forward the example above, let us suppose we are testing the following set of variables, all dichotomous for easy discussion here.

CONSTRUCTING AN INTERACTION TERM TO TEST WHETHER HIGH SCHOOL DROPOUT STATUS INFLUENCES POST-PROGRAM OUTCOMES LESS AMONG CT PARTICIPANTS THAN AMONG OTHERS.

VARIABLE	VALUES
PROGRAM ACTIVITY	(CT participants = 1; all others = 0)
EDUCATION	(dropouts = 0; HS grad or more = 1)
INTERACTION (PROGRAM ACTIVITY * EDUCATION)	(If CT and HS grad = 1; others = 0)

To test the interaction term, the interaction term, appropriate control variables, and the original variables from which the interaction was constructed must be included in cases like this, in which we have neither an experimental design nor a firm theory directing our interaction (Blalock, 1965). Since the interaction term will include portions of both original variables, an effect of either or both original variables would erroneously be carried by the interaction term alone. Only when the original terms are both included in the test can we be certain an effect of the interaction term is not spurious.

Let us assume that our OLS multiple regression results indicate the following slopes for the three variables of interest:

PROGRAM ACTIVITY	slope (b) = + .15
EDUCATION	slope (b) = + .30
INTERACTION	slope (b) = - .20

The hypothetical results above indicate a) that compared to all other treatments, CT raises post-program employment 15%, b) that high school graduation, in itself, raises post-program employment 30%, and c) that in addition to its positive effect on all clients, CT erases a portion of the effect of education, cutting back the advantage of high school graduation by 20%.

It should be noted that because interaction terms are by definition highly correlated with their original constitutent variables, careful analysts will consult the change in variance explained (R^2) rather than the t scores of the individual variable when deciding whether an interaction term in a regression model is statistically significant.

REPORTING STANDARDIZED OR UNSTANDARDIZED REGRESSION COEFFICIENTS

Standardized regression coefficients (Betas) are often reported because they indicate the relative power of each variable in one equation to account for variation in the dependent variable. Betas have a common sense meaning similar to that of a correlation; a Beta of .5 always indicates a "stronger" effect than a Beta of .4. Unstandardized coefficients (regression slopes) are expressed in terms of the metric of the independent and dependent variables, and are often less intuitively satisfying. If education is scored using a four point scale, a slope of .10 indicates that *each step* of that education scale raises the outcome variable by .10. The lowest step compared with the highest, three intervals above, has an estimated .30 higher level. If the dependent variable is employment status, .30 translates to 30%. If it is wage level, .30 translates to $.30.

These considerations make slopes somewhat more complex than Betas to communicate when findings are reported. However, when results of research are being applied to program development efforts, unstandardized slopes are often preferable to report. The reason is that they do not reflect the variation within a sample, but rather give a direct estimate of the *amount of change in an outcome which is produced by a given change in the input*. Is it more helpful for me to know that education, which I cannot control, is more powerful than program activity? Or am I better served by estimating that after the effects of education are accounted for, my OJT program produces post-program wages an estimated $.47 lower (or higher) than my CT program? Clearly the latter, unstandardized report is preferable.

In general, unstandardized regression slopes are both more useful and easier to report when the dependent variable is naturally interpretable, as in the case of income or a dichotomy which translates to percentages, and when the independent variable is a dichotomy, allowing statements like "Category A is X amount larger than category B."

When these conditions do not hold, the analyst must choose between reporting ease and managerial usefulness. For a full analysis, both forms augment each other. For example, a treatment experienced by only a few participants cannot account for much explained variance, and will therefore produce a low Beta, even if the effect of that variable is strong among the few to whom it applies. Therefore, the two reports might indicate both a negligible impact on overall outcomes, and a large impact on a few clients. When benefit-cost analysis is attempted, unstandardized slopes are required.

PART II.

ESTIMATING CHANGE

The first step in estimating change is, of course, measuring identical variables covering the pre-program and the post-program periods. (See Appendix E.) Analysis of these variables takes different forms for descriptive analysis and differential impact analysis.

For descriptive analysis, the pre-program value is subtracted from the post-program value, and the result reported as *change*. As with all descriptive analysis, the change figure is influenced by many unknowns. In the case of change, one of these is the original base figure. If, for example, a group of participants includes many with zero earnings during the pre-program year (e.g., students or displaced homemakers) then that group is likely to generate higher income change than a group including primarily high previous earners. This occurs simply because the change figure is based in part on the pre-program figure.

In multivariate analysis, the goal is to estimate unique causal effects of each factor tested. Therefore, the effect of pre-program rates on change must not be allowed to bias estimates of other variables on change. That is, two of the following three variables must be included in any multivariate test of change: 1) the *pre-program level,* 2) the *post- program level,* and 3) *change,* the post-program minus the pre-program levels. Since any two of these is sufficient to produce the third via simple mathematical operations, the equation is satisfied regardless of which two are used.

Therefore, it is much preferable to use the two which produce the most sensible results: the post-program outcome as dependent variable and the pre-program measure of the same outcome as a control variable. This latter variable is known as the autoregression term. It indicates stability over time, the tendency for pre-program patterns to reproduce themselves after the program. When it is included in an equation, other variables which show a reliable impact on an outcome may be interpreted as increasing or decreasing the rate of change from pre- to post-program levels of the outcome in question.

PART III.

ANALYZING A LARGE NUMBER OF POTENTIAL INFLUENCES ON OUTCOMES

Multivariate analysis is straightforward in studies where only a few theoretically derived variables are tested. All are entered into the equation and results reported. However, in cases where many measures are to be tested as independent variables, it is no longer possible or advisable to include all in a single test. Such attempts run up against degrees of freedom limitations, suffer from accumulated missing cases, or erode meaningful estimates because of the many small intercorrelations which occur by chance. This produces the problem of how to move through multiple tests efficiently and without distorting or overlooking effects. The following suggestions may assist in that endeavor, although they are of course not complete. Performing such analysis requires prior statistical background. This "cookbook" summary is not meant to imply otherwise. It only suggests steps to make analysis relatively efficient.

Step 1. Insure that variables are in the proper form for multivariate analysis and that variation is sufficient to make analysis meaningful.

Step 2. Select the appropriate dependent variable for each analysis.

Step 3. Separate variables according to their importance for test. Those which are most important, because they are known to affect the outcome and must be included to prevent bias or because they hold special program development interest should be given priority during analysis.

Step 4. Separate variables according to missing cases. In particular, questions asked only of subsets of participants, such as job qualities or reasons for unemployment, should be analyzed separately from other questions applying to all. The reason is that the safest form of multivariate analysis is based only on cases for which full information on all variables is present. Under that approach, any case with a single missing value is eliminated from the entire analysis. Therefore, variables with many missing values should be analyzed separately.

Step 5. Compute correlations between all independent variables and the dependent variable(s) being analyzed. Correlations are the basis for the calculation of OLS regression coefficients, making them the appropriate bivariate test building toward regression analysis.

Step 6. Identify those variables which are appropriate in terms of missing values and which have high priority as control variables -- those required to protect against biased estimates. Observe their correlations with the outcome(s) in question. Select those variables from this set which exhibit a reliable relationship with the outcome being analyzed.

Step 7. Enter the variables selected at the conclusion of step 6 into a multiple regression equation. Identify the subset of these variables which have reliable multivariate effects on the outcome. For simplicity, analysts may use a stepwise procedure, which automatically selects reliable effects. This produces a minimal set of control variables which must be included in subsequent runs. Other variables from these tests may be set aside for the moment with the knowledge that were they included, their effect would be too small to alter findings noticeably.

If the number of measures showing reliable correlations is large, more than one per 10-15 cases, divide them into smaller sets and repeat step 7 for each, combining the results into a final test. At any time that the number of cases included in an equation drops well below the total number in the sample, examine missing cases to determine why and correct the analysis.

Step 8. Identify the most important test variables -- program variants of special policy interest. Observe their correlations with the outcome in question, selecting those which show reliable association. Enter these singly or in appropriate sets into equations which include the minimal set of control variables identified in step 7.

Step 9. In addition to variables tested in step 8, analysts may wish to *explore* other program variants, hoping to discover useful unexpected relationships. Group measures according to policy area, such as intake, quality control, or trainer characteristics. Correlate these with the outcome being analyzed and enter those which are reliably greater than zero into an equation including the minimal set of control variables identified in step 7.

Step 10. The procedures outlined above for reducing the set of reliable effects ignores the possibility of suppression, the situation in which two variables are correlated with each other but have opposite effects on the dependent variables and therefore tend to cancel each other out in bivariate tests. These effects become visible only when both independent variables are tested jointly. They are therefore overlooked when only reliable bivariate correlations are forwarded for test in regression equations, as in steps 6-9.

Short of a full exposition of this issue, one step may be suggested which will guard against most errors of this type. Suppression of the type which the analyst most wishes to uncover occurs only when some variable is correlated with one of the variables identified as reliable during steps 8 and 9. Correlations should therefore be calculated between each of these reliable effects and other independent variables. Where reliable correlations are found, the variable in question may be added to the reduced set of reliable effects located after step 8 or 9. Relatively few changes will be produced by such a procedure, but it does guard against the most damaging errors from undetected relationships. These tests may be facilitated through use of a backwards stepwise elimination of unreliable effects.

Step 11. Membership variables such as service provider, Standard Industry Code category, training area, and the like should be examined. These may be tested as a series of dummy variables added to reduced sets of reliable effects (steps 8 or 9). Findings may prove useful for future contracting or marketing. Also, such tests protect against spurious findings of program effects, as discussed in Chapter 3.

Step 12. Having identified reduced sets of the most powerful and unique effects on each outcome being analyzed, the analyst will be well advised to return to the data set in order to examine what measures are associated with these key effects. Such analyses may be conducted formally, using these key effects as dependent variables in their own right, or may be less formalized examinations of patterned associations. The purposes of such further analysis are:

- To test associations which might corroborate interpretations formed on the basis of initial findings, or

- To develop interpretations of initial findings by examining the apparent nature of the variables found to have greatest impact on the outcome.

The general strategy of the steps listed above is to structure successive tests based on thoughtful interpretations of analysis goals and necessary controls, in such a way as to gradually reduce the number of variables included in the analysis. The goals can be either to test hypotheses guiding data collection or to zero in on the set of factors which exhibit greatest potential as guides to policy development.

Any one analysis is likely to err in some of its conclusions. That is particularly true of exploration such as suggested in step 9. If the rule for establishing statistical reliability is the 95% confidence level, then by chance alone about 5% of all tests will be shown to be reliable. They will be false positives. Although few of these are likely to survive successive multivariate tests, this fact cautions against overinterpreting findings of modest associations. It especially recommends placing most faith in interpretations based on several findings which reinforce each other.

PART IV

COMBINING DATA FROM DIFFERENT SOURCES

One complexity of effective differential impact analysis is the need to combine data from several different sources into a tailor made data set. The combination of data also provides two of the method's strengths: the ability to protect from several angles against selection bias, and the ability to measure different types of variables in the most reliable and efficient manner.

Aside from the availability of appropriate computer facilities, the key to combining (merging) data is the inclusion of the correct identifiers in each data set to be merged. In Appendix E, a sample *master identifier coversheet* is included immediately prior to the participant survey. This sheet is structured to include all the identifying information required to merge data.

For differential impact analysis, all data should be merged into individual participant records, since the participant is the unit of analysis. All identifiers must appear in the participant's original data file. Each other original file must contain only the particular identifier required to correctly merge into the participant file. For example, implementation program variants are measured at the agency level. Each participant who enrolled through the agency with the ID code "10" will receive values on all implementation variables which were provided by that agency. The agency identifier will appear on those participants' master identifier sheets and also on the appropriate agency implementation data reports, allowing the match of identifiers, followed by the combination of data.

Once data sets are merged, statistical tests will be calculated on the basis of the number of participants (or employers) in the data set, not on the basis of the number of service providers or geographical regions which may have supplied particular data elements. The analyst must therefore remain cognizant of limitations surrounding the number of separate treatment contexts required for reliable differential impact tests. (See Chapter 3.)

SECTION V.

APPENDICES

REFERENCES

REFERENCES

Babbie, Earl R., Survey Research Methods, 3rd ed., Belmont, CA: Wadsworth Publishing Company, 1984.

Barnes, Hilda N., "Finding and Interviewing the Hard-to-Locate: The DMI Experience," Chap. 11:pp 145-154, in Evaluating the Impact of Manpower Programs, M. Borus (Ed.), Lexington, MA: Lexington Books, 1972.

Bassi, Lauri J. "The effects of CETA on the postprogram earnings of participants," Journal of Human Resources, XVIII: 539-556, 1983.

Blalock, H.M., Causal Inferences in Nonexperimental Research, New York: W.W. Norton and Company, 1964.

Blalock, H.M., Causal Models in the Social Sciences, Chicago: Aldine, Atherton, Inc. 1971.

Blalock, H.M., "Theory building and the statistical concept of interaction," American Sociological Review, 30: 374-380, 1965.

Bloom, Howard S., "Estimating the effect of job-training programs, using longitudinal data: Ashenfelder's findings reconsidered." Journal of Human Resources, XIX: 545-556, 1984.

Bloom, Howard S., and M.A. McLaughlin, "CETA Training Programs -- Do They Work for Adults," Joint CBO-NCEP Report, 1982.

Borus, Michael E., "Assessing the Impact of Training Programs," Chap. 2: pp 25-40. in Employing the Unemployed, E. Ginzberg (Ed.) New York: Basic Books, Inc., 1980.

Borus, Michael E., Measuring the Impact of Employment-Related Social Programs, The W. E. Upjohn Institute for Employment Research, April, I979.

Borus, Michael and Tash, William, Measuring the Impact of Manpower Programs: A Primer, Ann Arbor, MI: Institute of Labor and Industrial Relations, The University of Michigan - Wayne State University, Nov. I979.

Bradburn, Norman M. "Response Effects", in P. Rossi, J.D. Wright, and A. B. Anderson (eds.) Handbook of Survey Research, New York: Academic Press, 1983.

Bryant, Edward and Rupp, K., "Evaluation of the Job Training Partnership Act: Implications of the CETA Experience", Westat, Inc., January I985.

Campbell, Donald T. and Cook, T., Quasi-Experimentation: Design and Analysis Issues for Field Settings. Chicago: Rand McNalley, 1979.

Campbell, Donald T. and Stanley, J. C., Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally, 1966.

Caporaso, James A. and Roos, L. L., Quasi-Experimental Approaches: Testing Theory and Evaluating Policy. Evanston: Northwestern University Press, 1973.

Cook, Robert, C. Resman, K. Rupp, W. Turnage, and associates, "Early Service Delivery Area Implementation of the Job Training Partnership Act, Westat, Inc. 1984.

Comptroller General's Report to the Congress: "Job Training Partnership Act: Initial Implementation of Program for Disadvantaged Youth And Adults", GAO/HRD-85-4, I985.

Corcoran, Mary and Hill, Martha S., "Reoccurrence of unemployment among adult men," The Journal of Human Resources, XX:2, pp 165-183,1985.

Danziger, Sandra, "Postprogram changes in the lives of AFDC supported work participants: a qualitative assessment," The Journal of Human Resources, XVI:4, pp 637-648, 1981.

Dillman, Don A. Mail and Telephone Surveys: The Total Design Method. New York: John Wiley & Sons, 1978.

Director, Steven and Schneider, Stephan A., "Short-Term Evaluation and Research Issues for the Job Training Partnership Act of l982--Impact of Different Forms Of Private Sector Participation on Client Characteristics and Program Outcomes: Final Report," Research Report Series, National Commission for Employment Policy, Washington, D.C., December l983.

Doeringer, Peter B and Piore, Michael J., Internal Labor Markets and Manpower Analysis. Lexington, MA: D.C. Health, l971.

Franklin, Grace and Ripley, Randall B., CETA: Politics and Policy, l973-l982, Knoxville, TN: The University of Tennessee Press, 1984.

Gay, Robert S. and Borus, Michael, "Validating performance indicators for employment and training programs," The Journal of Human Resources, XV:1, pp. 29-48, l980.

Geraci, Vincent J., "Short-Term Indicators of Job Training Program Effects on Long-Term Participant Earnings", U. S. Department of Labor, August, l984.

Gillespi, M. W. "Log-linear techniques and the regression analysis of dummy dependent variables: further bases for comparison," Sociological Methods and Research, 6:103-122, 1977.

Goodman, L.A., "A Modified multiple regression approach to the analysis of dichotomous variables, " American Sociological Review, 37:28-46.

Grembowski, David, "A Guide for SDA Process Evaluation", Olympia, Wa.: published for the JTPA Evaluation Design Project by Washington State Employment Security Department, 1986a.

---------------- "A Guide for State Process Evaluation", Olympia, Wa.: published for the JTPA Evaluation Design Project by Washington State Employment Security Department, 1986a.

Gurin, G. , A National Attitude Study of Trainees in MDTA Institutional Programs. Ann Arbor, Michigan: Institute for Social Research, l970.

Homans, Celia, "Finding the Hard-to-Locate: The NORC Experience," Chap. 12: pp. 155-164 in Evaluating the Impact of Manpower Programs, M. Borus (Ed), Lexington, MA: Lexington Books, 1972.

Johnson, Terry R. "A Guide for Net Impact Evaluation", Olympia, Wa.: published for the JTPA Evaluation Design Project by Washington State Employment Security Department, 1986.

Johnson, Terry R, Dickinson, Katherine P., West, Richard W., "An evaluation of the impact of ES referrals on applicant earnings," The Journal of Human Resources, XX:1, pp117-136, 1985.

Kiefer, Nicholas M., "The Economic Benefits From Four Government Training Programs," pp159-186 in Evaluating Manpower Training Programs, F. Block (Ed.), Greenwich, Conn.: JAI Press Inc., 1979.

Kish, Leslie, Survey Sampling. New York: John Wiley & Sons, 1965,

Knoke, D. "A comparison of log-linear and regression models for systems of dichotomous variables," Sociological Methods and Research, 3: 416-434, 1975.

Kobrak, Peter, Private Assumption of Public Responsibilities: The Role of American Business in Urban Manpower Programs. New York: Praeger, 1973.

Lewis, Morgan V., "Finding the Hard-to-Locate: A Review of Best Practices," Chap. 13: pp. 165-173, in Evaluating the Impact of Manpower Programs, M. Borus (Ed.), Lexington, MA: Lexington Books, I972.

Levitan, Sar A., and Magnum, Garth L., "The T in CETA: Local and National Perrspectives", Kalamazoo, MI: W. E. Upjohn Institute for Employment Research, I981.

Mangum, Garth and Robson, R. T., Metropolitan Impact of Manpower Programs: a Four-City Comparison, Salt Lake City, UT: Olympus Pub., I973.

Maranto, Cheryl L. and Rodgers, Robert C., "Does work experience increase productivity? A test of the on-the-job training hypothesis", The Journal of Human Resources, XIX:3, pp 341-357, 1984.

Minnesota Office of Statewide CETA Coordination, "Private Sector Experience with CETA On-the -Job Training", St. Paul, MN, Minnesota Department of Economic Security, August, I979.

Mirengoff, W. and Rindler, L. The Comprehensive Employement and Training Act: Impact on People, Places, Programs. National Academy of Sciences, 1976.

National Academy of Sciences, "High Schools and the Changing Workplace: The Employers' View", Washington, D.C.: National Academy Press, , I984.

National Alliance of Business, "What's Happening with JTPA? A Complete Analysis of NAB's 1984 Survey Data", Washington D.C., National Alliance of Business, 1985

National Commission for Employment Policy, "CETA Training Programs--Do They Work for Adults?" Congress of the United States, Congressional Budget Office, July I982.

National Governor's Association, "Assessment of Adult and Youth Performance Standards Under the Job Training Partnership Act," 1985.

O'Neill, Dave, "Employment tax creditprograms: The effects of socioeconomic targeting provisions," The Journal of Human Resources, XVII:3, pp. 449-459, 1982.

Perry, C.R., Anderson, B., Rowan, R. L., and Northrup, H., The Impact of Government Manpower Programs. Pennsylvania: Industrial Research Unit, The Wharton School, U of Penn, I975.

Rossi, Peter H. and Freeman, Howard E., Evaluation: A Systematic Approach. Beverly Hills, CA.: Sage Publications, 1982.

Rossi, Peter H., Wright, J.D. and Anderson, A. B., Handbook of Survey Research, New York: Academic Press, 1983.

Sawhney, Pawan K., Jantzen, Robert Herrmstadt, Irwin L., "The differential impact of CETA training". Industrial and Labor Relations Review, Vol. 35 No. 2 (January I982) by Cornell University.
Schiller, B.S., "Lessons from WIN: a manpower evaluation," The Journal of Human Resources, XIII:4, pp. 502-523, 1978.

Seattle-King County Private Industry Council and Snedeker Scientific, Inc. "JTPA Evaluation Issues, Priorities, and Contingencies at the SDA Level" Technical Report for the JTPA Evaluation Design Project, Washington State Employment Security Department, 1985.

Simpson, Carl, "CETA On-the-Job Training Study Final Report", Olympia, Wa. : Washington State Employment Security Department, 1984a

------------ "CETA On-the-Job Training Employers: A Survey of Employers and OJT Officers", Olympia, Wa.: Washington State Employment Security Department, 1984b.

----------- "Vocational Classroom Training Outcomes Project: Final Report" Olympia, Wa.: Washington State Employment Security Department, 1982.

Snedeker, B and Snedecker, D. CETA: Decentralization on Trial. Salt Lake City, UT: Olympus Publishing, 1973.

Streker-Seeborg, Irmtraud, Seeborg, Michael C., Zegeye, Abera, "The impact of nontraditional training on the occupational attainment of women," The Journal of Human Resources, XIX:4, pp 452-471, 1984.

Stromsdorfer, Ernst, "Issues in Evaluating Costs and Benefits," Olympia, Wa.: published for the JTPA Evaluation Design Project by Washington State Employment Security Department, 1986a.

Sudman, Seymour, Applied Sampling, New York: Academic Press, 1976.

Sudman, Seymour and Bradburn, N., Asking Questions, San Francisco: Jossey-Bass Publishers, 1982.

Taggart, Robert, A Fisherman's Guide: An Assessment of Training and Remediation Strategies. Kalamazoo, Mich: W. E. Upjohn Institute for Employment Research, 1981.

U.S. General Accounting Office, "CETA Programs for Disadvantaged Adults--What Do We Know About Their Enrollees, Services, and Effectiveness?" Report to the Chairman, Subcommittee on Employment Opportunities, Committee on Education and Labor United States House of Representatives, June 14, 1982.

Vermeulen, Bruce and Hudson-Wilson, Susan, "The Impact of Workplace Practices on Education and Training Policy," Chap 4, pp 72-87 in Jobs and Training in the 1980s: Vocational Policy and Labor Market, Peter Doeringer and Bruce Vermeulen (Eds), Boston: Martinus Nijhoff Publishing, 1981.

Walker, Gary, H. , et al.," An Independent Sector Assessment of the Job Training Partnership Act, Phase I: the Initial Transition", Grinker Walker, Inc. and others, 1984.

Walker, Gary, H. Feldstein, and K. Solow," An Independent Sector Assessment of the Job Training Partnership Act, Phase II: Initial Implementation", Grinker Walker, Inc. and others, 1985.

Weinberg, Eve, "Data Collection: Planning and Management", in P. Rossi, J.D. Wright, and A. B. Anderson (eds.) Handbook of Survey Research, New York: Academic Press, 1983.

Weiss, Carol H., "Interviewing in Evaluation Research," pp. 355-395 in Handbook of Evaluation Research, E. Struening and M. Guttentag (Eds.), Beverly Hills, CA.: SAGE Publications, 1975.

Wentling, Tim and Lawson, T., Evaluating Occupational Education and Training Programs, Boston, MA.: Allyn and Bacon, Inc., 1975.

West, Richard and Dickinson, Katherine, "Development of Adjustment Models for PY 85 JTPA Performance Standards," Menlo Park, CA, SRI International, April 1985.

West, Richard and Dickinson, Katherine, "Issues Involved in Sampling to Obtain Postprogram Measures of Performance, Menlo Park," CA, SRI International, June 1985.

Westat, Inc., Summary of Net Impact Results. Rockville, MD: Westat, Inc., 1984.

Zornitsky, Jeffrey, "Major Issues in Designing a Post-Program Performance Management System for the Title IIA and Title III Programs Funded Under JTPA," Cambridge, MA, Abt Associates, Inc., March, 1985a.

Zornitsky, Jeffrey, Fairchild, C., Jastrzab, J. and Seitchik, A. "Measuring the Costs of JTPA Program Participation," Cambridge, MA: Abt Associates, Inc., Sept. 1984.

Zornitsky, J., Schneider, G., Sharick, and Sharpiro, L., "Post-Program Performance Measures for Adult Programs Funded under Titles IIA and III aof JTPA: Early Findings Prepared for the U.S. Department of Labor's Performance Standards Advisory Committee," Cambridge, MA : ABT Associates, Inc., June 1985a.

APPENDIX A

Appendix A. Review of Selected Literature

In the case of the gross impact approach, a wide array of previous work touches on topics of interest, but few illustrate the entire approach. It is therefore not possible to trace a single line of past work showing the origins and development of the gross impact approach. Indeed, one of the major potentials of this approach is its promise of producing new types of information for JTPA policy makers.

STUDIES OF PROGRAM IMPACT

The most fully developed literature in the field of employment training evaluation is literature on the net impact of training programs. (See Johnson, 1986 for a thorough review). Although the thrust of that work is quite different from that of gross impact analysis, it also provides findings which are useful background to gross impact analysis.

A succinct summary of studies published prior to the availability of the CETA Continuous Longitudinal Manpower Surveys (CLMS) is offered by Borus (1980). He concludes from 25 published research estimates that training does have modest positive effects, that classroom training has a relatively consistent impact of raising income $300-400 (in 1980 dollars), that the on-the-job training impact averaged about twice that amount but that different estimates varied widely, that youth work experience raised incomes but that estimates vary widely among the studies available, and that the few studies of adult basic education and adult work experience showed modest positive program impact on earning.

A more recent set of estimates reflect improved methods and data bases, although most derive from one source: the CLMS. These are well reviewed in three separate sources, all of which became available in 1981-1982 (Bloom, 1982; Taggart, 1981; National Commission for Employment Policy, 1982). Each of these is highly readable, written for a relatively wide audience. Despite their shared data base, this new set of estimates shows considerable variation depending on the assumptions made and on the method of analysis, making any summary hazardous. Nonetheless, a broad summary probably includes the points below.

The overall CETA program probably had a very slight positive impact on earnings, with most estimates between zero and $300 per year. The impact on men is estimated to be small or negative, while the estimated impact on women is consistently positive. Some studies estimate large gains for women, with gains for black women smaller than for other women. Adult work experience is the program showing poorest, usually negative, impact. Both pre-program and post-program earnings are higher for OJT than for CT, with their relative net impact estimates depending on method of analysis. Adding a recent study of ES referrals as roughly equivalent to JSA, we can estimate that the net impact of JSA is similar to that for other interventions: no impact among males; substantial positive impact among females (Johnson, 1985).

The predominance of net impact studies leaves the gross impact approach with both the advantages and the disadvantages of research in a relatively undeveloped area. The disadvantages are:

1. That the method remains unrefined, although we receive some help from the field of higher education where this method has been widely used, and

2. That few previous research findings are available to guide our choice of specific measures or hypotheses.

The advantage is that even the most basic of findings using this approach have the potential to offer new and useful insights to JTPA policy- makers. Net impact research has focused almost exclusively on wage gain or loss, leaving many program outcomes unexamined. National studies have tended not to address state and local program issues, and net impact studies' focus on overall training programs or on their major variants has left unexamined the question of what differential impact more specific local program variants may have.

EARLY JTPA IMPLEMENTATION STUDIES

A second research tradition within the employment and training field is less precisely developed than the net impact tradition, but probably reaches an even larger audience. This is implementation research, examining how programs are formed, how legislative mandates are met, what organizational patterns emerge, and what problems appear to predominate at particular points in program history. The best of these studies offer numerous concrete insights for program managers as well as for national policy makers (Mirengoff and Rindler, 1976; Snedeker and Snedeker, 1973; Levitan and Mangum,1981).

At this writing, several major JTPA implementation studies are available, providing a view of the major policies and approaches being followed by most SDAs. (National Alliance of Business, 1985; Walker, 1984, 1985; Cook, 1984; National Commission for Employment Policy, 1982). These are thoroughly reviewed by Grembowski (1986a; 1986b). Here, a few points concerning the issues of basic SDA program policies and service level are highlighted. One component of the gross impact approach is the systematic measurement of selected program implementation or individual treatment process variables, in order to estimate their relative effectiveness in producing desired outcomes. These issues raised by JTPA implementation studies offer some guidance as to process issues most uniquely relevant to JTPA.

All these implementation studies emphasize the potential conflict JTPA regulations cause between serving those most in need and producing high performance ratings, including low costs and high placement rates. Walker (1985) finds a perhaps alarming consistency in SDA decisions to serve the local employment market's needs rather than the needs of the hardest to serve. Numerous SDA directors offer quotations stating in essence that to fulfill the new legislation's intent, they should maximize performance and employer benefits in opposition to participant benefits. The political reasoning is made clear by one such quotation: "This legislation won't tolerate failure. That means it won't tolerate risk." (Walker, 1984: 78) Serving those with genuine need constitutes risk.

The Cook and National Alliance of Business (1985) studies suggest somewhat greater balance between the needs of employers, performance statistics, and participants. However, Cook notes that the two major routes to retaining service quality represent compromises to insure performance levels. In one adaptation, short and inexpensive treatments for the most job-ready insure high placement while subsidizing the costs of the rest of the program geared to those in need. In the second, the bulk of the program serves the relatively job-ready, allowing a few small programs for those most likely to benefit from intensive training efforts.

The nature of early JTPA program performance illustrates this issue: most SDAs met entered-employment and cost per entered-employment standards, but the average wage standard was met much less often. That is, it appears that quantity is being gained at least in part through low quality placements. Since all traditional justifications for federal intervention in this arena have rested on some assertion of client need for assistance and the consequent probability that clients' lives would be changed by intervention, these concerns about service quality cut to the heart of training program philosophy (Taggart, 1981). As such, they force the attention of future research efforts toward issues of service quality and the question of how outcomes are accomplished. Net impact research is necessary to determine the nature of overall program impact. However, process and gross impact research, measuring a fuller range of variables, is necessary to assess the nature of the service or service quality.

The other side of the service quality coin is that classroom skills training and on-the-job training are becoming the dominant programs. Since net impact studies consistently show these programs to be preferable to work experience, and since other options tend to have ambiguous relationships to employment, this may be seen as a high quality program mix. That assessment is tempered, however, by the predominance of very short programs.

A second type of finding which emerges from JTPA implementation studies identifies some major policy dividing points among SDAs. These are identified as basic variations within allowable policy tolerances, with a considerable number of SDAs following each approach. As such, they are especially interesting for comparative research. These include the following:

1. Orientation to serve
   - those most job-ready
   - those most likely to benefit from training, or
   - those most in need of assistance (Cook, 1984).
2. Orientation to participant needs, to employer benefits, or to employer marketing attempts (Walker, 1985).
3. Targeting practices: including
   - Only as required,
   - Requirements plus other groups (most frequently high school dropouts and minorities (National Alliance of Business, 1985), or
   - Enhanced targeting reinforced by added formal state performance requirements.
4. The decision to provide stipends for classroom training or not to do so: about half of SDAs use some, with the typical allowance being relatively small and restricted to those with greatest need (National Alliance of Business, 1985).
5. Centralization of intake versus intake performed by service providers, as was typical under CETA.
6. Policy regarding enrollment in multiple services. The National Alliance of Business (1985) reported that SDAs plan for about one-third of all participants to enroll in more than one service.

In addition to these service delivery policies, another issue raised by implementation studies holds particular interest for gross impact comparisons: the types of internal record keeping and analysis performed by SDAs. Cook (1985) reports that half of the SDAs studied did not keep data on termination statistics by program. That is, they did not perform even the most basic of systematic internal diagnosis. However, interest in internal analysis appears to be growing. State level interviews by the National Governors' Association (1985) found 29 states keeping records on characteristics of all eligibles, rather than enrollees only, and fully 45 states conducting or planning some type of follow-up data collection effort.

More recent interviews found most SDAs highly concerned about performance accountability, with many producing monthly reports calculating performances by program. (Seattle/ King County, 1985). In addition, they found considerable interest in expanding use of evaluation analyses. Major areas of development include:

1. beginning longer term follow-ups or enriching follow-ups with employers or participants,
2. performing more detailed analyses of follow-up data,
3. upgrading MISs,
4. conducting process evaluations or special assessments,
5. creating new data linkages (e.g., with UI or welfare systems).

These JTPA implementation findings provide a useful set of directions to follow when deciding which aspects of JTPA programs and administration to include in differential impact comparisons of service delivery alternatives.

DIFFERENTIAL IMPACT ANALYSIS

Formal, quantitative differential impact analysis of program variants is a relatively underdeveloped field, as is well illustrated in Borus's (1979) program evaluation primer. After listing 44 specific participant characteristics known to affect labor market success, he turns to the question of "program component independent variables." His one paragraph discussion of this topic begins by saying that "It would be extremely useful in modifying existing programs and in the planning of new programs to know which of the (program) components are most effective for various types of participants" (p 70). Two measures are suggested: program length, "... and, if possible, a measure of quality." This paucity of guidance is not pointed out to fault Borus. On the contrary, his work provides an excellent example of this point because it is so highly respected.

Some early studies compared several distinct programs such as MDTA, Job Corps, and WIN (Mangum & Robson, 1973; Keifer, 1979). While these involve estimates of differential impact, the options are not local - they are national legislative funding options. In extensive 1982 summaries of CETA impact studies (National Commission for Employment Policy,1982; Taggart, 1981) the only differential comparisons included were comparisons among basic program options, AWE, OJT, CT, and PSE, and findings concerning length of exposure to the CETA program. These are differential comparisons in the sense used in this gross impact guide. They are minimal, yet useful, since basic program mix is one of the SDAs' most effective tools for improving program outcome levels.

The relatively undeveloped state of differential impact analysis is illustrated by findings concerning length of participation. Higher entered-employment and earnings occurred where CETA participation was longer. However, most of these analyses confound four very different phenomena: intended program length, program completion versus early withdrawal, extended program time through part-time study or remedial extensions, and up to 90 days of post-program job search time prior to formal termination. Taggart (1981) untangles one of these factors, non-completion of the program, and finds a large effect of non-completion along with a small effect of the remaining variation in participation time. To fully untangle these factors requires much more specific measurement.

Schiller (1978) gives us an example of differential impact analysis by comparing five different levels of WIN intervention: no services, job search services, education, training, and subsidized employment. He finds program impact "...highly sensitive to the mix of services actually provided" (p 513). In particular, impact was much more positive for subsidized employment than for training, and training more positive than other options. In this instance, impact followed cost, making impact alone an inadequate base for decision making, but also making cost alone an inadequate base.

Another example is provided by Streker-Seborg et al. (1984), who analyze CETA training for women in non-traditional fields. They find women who trained in non-traditional fields to be employed as often as those who trained in other fields, but not to enjoy the greater income typical for men in those fields. The difference between these findings and those of Simpson (1982) offer opportunity to illustrate the relatively local flavor likely to characterize early differential impact studies, since we are just beginning to chart the mechanisms through which programs have their impact. In Simpson, women who took classroom training in non-traditional fields found work less often, but earned more when they found work than did women in traditional fields. In the case of OJT, the findings more closely paralleled those of Streker-Seborg.

Two currently published works devote extensive consideration to differential analysis of local program options: Levitan and Mangum (1981) and Franklin and Ripley (1984). They compare local program management and planning among small numbers of CETA prime sponsors, analyzing, among other things, their likely differential impact on program outcomes. Most of their analysis was performed at the aggregate level, with small samples, and Franklin and Ripley do not present specific findings in their text, leaving some questions about the reliability of conclusions. However, these works point in valuable directions. Some of the questions they ask and the tentative answers they receive illustrate this.

Probably the most important observation to make about these studies is that they observe and compare matters over which state and local policy makers have considerable control, such as management practices, staff qualifications, and program mix, and find relatively strong impact on placement rates. Such findings may "confirm the obvious" for program managers, who would presumably have difficulty doing their jobs without a faith that local management decisions make a difference. Yet they are relatively unique in the employment and training literature. For example, use of requests for proposals (RFPs) is related to lower cost and higher placement rates, as is staff quality, explicit staff emphasis on placement as a goal, high use of OJT, and low use of WEX (Franklin and Ripley, 1984: 92). In addition, both authors report that these manipulable factors have greater impact on outcomes than externally set factors such as unemployment rates and client characteristics.

Simpson's (1982; 1984a) recent work in Washington state also approximates the differential analysis model. It is not reviewed in any detail here since extensive use is made of examples from that research in Appendix C. Suffice it to say here that two CETA studies were undertaken, one of CT and the other of OJT, in which a wide range of program process variables were measured by interviewing prime sponsor staff and individual participants. Ability to explain program outcomes using these variables was limited, but program variants explained more of the variation than did individual participant characteristics.

BROADENING THE RANGE OF OUTCOME MEASURES

One reason the gross impact model is included in the JTPA Evaluation Design Project is to provide a mechanism for the collection of original post-program data from participants and employers, thereby allowing a wider range of outcome variables to be measured. Numerous measures have been suggested in past guides for evaluating job training programs, although the great bulk of studies have analyzed only a few. A particularly concise list was developed by Borus (1979, p 21-25), including benefits for society and the following benefits for individual participants:

1. increased incomes,
2. reduced unemployment,
3. increased satisfaction with work and with general conditions,
4. increased social status,
5. increased voluntary leisure,
6. reduced dependency,
7. improved health,
8. improved family life, and
9. improved housing.

The earnings and employment variables may be broken into more specific measures such as number or percent of hours or weeks worked and hourly wage rate. These in turn can be measured at the point of follow-up, cumulatively during the period between termination and follow-up, or as change from pre-program rates (Zornitsky, 1985a; 1985b).

To this list may be added several different types of outcome measures offering insight into how programs produce their effects. One is whether skill developed during the program is used in the placement position, among those employed. This outcome is valuable because it is by definition based on the program intervention. It is also correlated with various job qualities which indicate placement in the primary job market and is sensitive to both participant orientation toward skill use and program training variables (Gurin, 1970; Simpson, 1982).

Noneconomic outcomes are recognized in program evaluation, but seldom measured. After reviewing more than two hundred studies, Perry, Anderson, Rowan, and Northrup (1975) specify several types of post-program noneconomic outcomes of interest, including job satisfaction, better housing, greater leisure, greater sense of security (all based on economic variables), sociability, self-esteem, confidence, increased formal education, and enhanced citizenship. They reach three conclusions regarding this set of measures. First, only a few studies, 17 of 200 reviewed, include any information on impacts other than employment status and earnings. None of these was longitudinal, and none involved analysis any more sophisticated than reporting overall satisfaction scores or perceptions. Second, the literature which is available offers scant guidance as to whether training programs in fact have any impact on these outcomes. Third, based on interviews with Washington D.C. administrative officers of eight different national training programs, they conclude that: "The paucity of data on the noneconomic effects of manpower training does not allay the widespread belief among many program administrators that such benefits are pervasive and worthwhile." (p 96). It may be that the most cogent reason for testing such outcomes is as a validity check on such perceptions, either to encourage the recognition of such outcomes or to prevent self-deception as to unseen program benefits.

Danziger (1981) adds another dimension to the question of identifying meaningful outcomes by allowing participants to define outcomes most meaningful to themselves. This occurred during in-depth interviews at one year follow-up with Supported Work program participants. "When asked how they had been getting along since Supported Work, respondents described:

1. their employment status,

2. their family's financial situation, and

3. their own self-image, personal feelings of satisfaction and competence." (p 640)

All three types of reports appear to vary with 1. primary labor market employment versus continued welfare dependency and 2. personal drive or sense of efficacy.

Zornitsky et al. (1985a) attempted to identify post-program measures which are the most valid indicators of longer term net impact on earnings gain, calculated under varying assumptions. They summarize their analysis as follows: "Of the eleven measures tested, four consistently appear in the top five when assessed across the three bench marks. These include earnings, hourly wages at both the end of the post-program period and during the period, and percent of hours worked. They also note that follow-up measures are better predictors of long term gain than termination measures and that six month follow-up measures are more reliable as predictors than are three month follow-up figures.

Two broad conclusions may be drawn from this brief discussion of measures. First, although there is a wide range of possible outcomes open to impact research, little is known about most of these, and what is known points toward measures of employment and earnings status. Second, among measures of employment and earnings, there appears to be considerable utility in collecting data on measures not available through UI wage reports. Wage levels, hours worked, and the training-relatedness of employment and measures of job quality all add significant power to the analysis of program impact.

EMPLOYER BENEFITS

Consistent with the gross impact guide's role as the location for methods relatively new to the employment and training field, another goal of the guide is to encourage systematic employer follow-up surveys. Employer follow-ups, like client follow-up surveys, are emerging rapidly from the "SDA grass roots," (e.g., Seattle-King County,1985) but are so new to the formal research literature that extremely little published material is available for guidance. Washington state follow-up interviews with 600 OJT employers will be used extensively for illustrations of this approach (Simpson, 1984b).

Borus (1979) proposes three employer benefits, directed toward employers hiring the products of training programs, rather than engaging in subsidized hires as part of programs. These are:

1. Filling jobs of specific employers, such as bottleneck industries with hard-to-fill positions,

2. Filling jobs in particular geographical areas experiencing labor shortages, and

3. Improved productivity of an employer's labor force, such as increased labor output.

Although the third is the most difficult to measure, it is also the only one of the three likely to have meaning in periods of high unemployment.

Taking another track, a report by the National Academy of Sciences (1984) begins by listing competencies employers are likely to demand of high school graduates in the future. While JTPA participants are often not high school graduates, this list reflects the demands placed on primary labor market positions, and therefore represents program goals. Using this beginning, one might inquire as to how much JTPA training brings participants to minimal performance levels on such dimensions. However, the list covers a large range of skills from specific work habits to command of the English language. Since it is unlikely that 12 weeks of JTPA training will substitute for 12 years of school, this approach holds little specific value for employer surveys. Even so, the idea that one might ask employers to rate JTPA hires on several dimensions relevant to hiring policies is a useful one.

Vermeulen and Hudson-Wilson (1981) offer a useful back door approach to employer benefits by analyzing how employers avoid the costs associated with turnover. These costs include hiring costs, such as advertising, screening applications, interviewing, and testing, and costs of introductory training. Employers can reduce these costs by reducing turnover, which in turn may be accomplished by increasing pay to attract and hold better workers, by investing more heavily in recruitment to locate the best workers, or by lowering standards so that each hire, while less likely to last, costs less to accomplish. JTPA offers other avenues to reduce costs from turnover. Classroom training may enrich the local labor pool, making it easier to find skilled, and career-committed labor at lower wages. OJT offers multiple benefits, subsidizing screening costs, providing specifically trained workers at reduced cost to the employer, and providing a subsidized period to assess whether the participant is a desirable long term hire. Finally, brief try-out programs can be effective in reducing screening costs, but usually only in low-training jobs.

The two fundamentally different employer strategies for reducing costs of turnover identified above - increasing worker quality or decreasing job quality - have very different implications for employer benefits from JTPA. In what has come to be called the primary labor market, employers invest in high quality workers and emphasize retention and an internal promotion ladder. In these cases, training interventions are likely to produce benefits consistent with employers' priorities. In the secondary labor market, however, employers typically make little investment in training or job quality, and hire relatively young, untrained persons or those otherwise unable to demand high wages or high quality work. This creates a high turnover due to voluntary quits, but keeps the cost of turnover low. In these cases, serious training is unlikely to be viewed as an important benefit. Employers will tend to view a subsidy to hire persons very much like those normally hired a considerable benefit, although such hires represent poor program service to participants. This serves to remind us that any measures of the JTPA participants hired by an employer must be interpreted in light of the qualifications of persons who would otherwise have been hired.

Finally, measuring employer benefits also provides opportunity to collect employers' suggestions for improving programs and information useful for future marketing. Writing about evaluations of vocational education, Wentling and Lawson (1975) suggest the following objectives for an employer survey: assess performance of participants, compare program graduates with others, elicit employer recommendations for program improvements, determine recruitment practices of local employers, estimate local supply and demand patterns, and do public relations work for the program (see also Minnesota, 1979).

APPENDIX B

Appendix B. Background to Gross Impact Research Methods: Selected Topics

This appendix is segmented. It contains several discussions which are related to each other only in that they pertain to the methods suggested in this guide. Each provides some background or elaboration placed here in order to shorten the main document.

SAMPLE SIZE AND SAMPLING ERROR

As a supplement to Chapter 4's discussion of sample size, Exhibit B.1 shows the estimated sampling error assuming particular sample sizes and particular measures. It shows estimated error for dichotomous measures in which the proportion of responses falling into each of the two answer categories varies from 50%--50% to 90%--10%. These calculations offer analysts an approximate range of error likely to occur with samples of different sizes. The figures shown in Exhibit B.1 are estimated 95% confidence interval sampling errors. That is, these figures represent two times the standard error of estimates which would be generated under each set of conditions specified.

Exhibit B.1 Estimated Sampling Error

Sample Size	Percentage Distribution				
	50/50	60/40	70/30	80/20	90/10
100	10	9.8	9.2	8	6
200	7.1	6.9	6.5	5.7	4.2
300	5.8	5.7	5.3	4.6	3.5
400	5	4.9	4.6	4	3
500	4.5	4.4	4.1	3.6	2.7
600	4.1	4	3.7	3.3	2.4
700	3.8	3.7	3.5	3	2.3
800	3.5	3.5	3.2	2.8	2.1
900	3.3	3.3	3.1	2.7	2
1000	3.2	3.1	2.9	2.5	1.9
1100	3	3	2.8	2.4	1.8
1200	2.9	2.8	2.6	2.3	1.7
1300	2.8	2.7	2.5	2.2	1.7
1400	2.7	2.6	2.4	2.1	1.6
1500	2.6	2.5	2.4	2.1	1.5
1600	2.5	2.4	2.3	2	1.5
1700	2.4	2.4	2.2	1.9	1.5
1800	2.4	2.3	2.2	1.9	1.4
1900	2.3	2.2	2.1	1.8	1.4
2000	2.2	2.2	2	1.8	1.3

How to use this table: Find the intersection between the sample size and the approximate percentage distribution of the variable in question. The number appearing at this intersection represents the estimated sampling error, at the 95 percent confidence level, expressed in percentage points (plus or minus).

COMPARING THE THREE MAJOR TYPES OF SURVEYS

There is essentially no question that gross impact participant follow- up data should be gathered through original telephone interviews. Original data are demanded because of the concern for quick turnaround and definable follow-up period, and because of the model's orientation toward measuring a range of outcome variables. Among the forms of original data collection, telephone surveys are clearly superior to in-person interviews or mail interviews in this particular case.

In-person interviews are especially useful when the interview is very long or complex, or when the interviewer needs to be able to observe facial reactions or the respondents' living situation. None of these applies here. In-person interviews are therefore too expensive to seriously consider. Weinberg (1983) estimates that 5 hours of interviewer time must be budgeted for 1 hour of face-to-face interview, because of travel time and rapport-setting time. Given that gross impact follow-up interviews can remain well under 15 minutes, the added 4+ hours would increase interviewer time by 16 fold! In addition, face-to-face interviews require much more careful selection and training of interviewers, are more subject to response bias (especially social desirability bias), and have recently been experiencing increasing refusal rates (Dillman, 1978). Happily, earlier concerns that telephone interviews would be less reliable than face-to-face interviews have proven unfounded. (Dillman, 1978; Weiss, 1975; Weinberg, 1983; Bradburn, 1983; Babbie, 1984).

The other major option to telephone surveys, the mail questionnaire, is less expensive, but nevertheless not to be recommended for gross impact analysis. The primary drawback to mail surveys is their relatively low response rate. While response rates can be raised through careful preparation and reminder mailings such as with Dillman's "total design method," the potential for biased results remains high. For example, Dillman's great success with mail surveys, averaging 74% completion among a wide range of studies, is contrasted to his response rates averaging over 91% using telephone surveys. Further, response to mail questionnaires is especially low among less educated, disadvantaged populations. It is also important that a high proportion of mail survey non-response is caused by the refusal to participate. This introduces the possibility of serious bias on the basis of positive versus negative feelings toward JTPA. With phone surveys, non-response stems almost entirely from failure to locate individuals rather than from refusal.

Mail questionnaires also suffer reliability problems because some proportion of respondents experience confusion over questions or instructions. These can be clarified over the phone, but not through the mail, increasing both the amount of missing data and the amount of questionable data generated by mail questionnaires.

Set against these advantages of phone interviews is the issue of cost. Dillman (1978) places costs of mail surveys at under half that for phone surveys. However, about half the cost of these national phone surveys was for telephone toll charges. This means that it is possible for local projects with slim budgets to reduce the costs of telephone surveys. In addition, one cost of mail surveys is difficult to quantify and is therefore underrepresented in reported cost comparisons: the need for repeated mailings, handling, and record keeping places heavy demands on clerical time and reliability, and causes considerable delays in data collection. Thus, sheer nuisance value makes this form of data collection problematic to organizations other than full time research organizations. Phone surveys require many fewer discrete organizational steps and are therefore possible to conduct locally, in order to reduce costs.

Finally, of course, any analysis projects planning to combine data collection with required performance standards are required to use telephone surveys.

HOW VALID AND RELIABLE ARE SURVEYS?

There is an extensive literature on this subject which should not be reviewed in detail here. In general, this literature offers consistent reason for faith in survey methods. Many potential biases have been examined; only a few appear to be problematic, and none deadly. Here, only issues identified as potentially problematic for gross impact analysis are reviewed briefly.

Reliability

Reliability is always an issue. Research is reliable if it would reach the same conclusions when repeated. The conclusions might be right or wrong; they are reliable to the extent that random noise does not distort conclusions. For example, measures such as sex and age are reliable because they are unambiguous, known to the respondent, and easy to record accurately. No research is entirely reliable, but surveys score well in this arena (Babbie, 1984). One type of reliability problem is eliminated by surveys better than by most other research modes: by carefully constructing answer categories and computer coding, error during the recording of answers can be minimized. Use of pretested "pre-coded" answer categories also improves reliability by clarifying the meaning of the question and the nature of the answers which can be entered as data. Surveys are strong in these areas, given that they are constructed well.

However, the apparent simplicity of "just asking some questions" is deceptive; surveys' reliability is based on careful construction. Sources such as Sudman and Bradburn (1982), Dillman (1978), and Babbie (1984) offer guidance, but local attempts to write original surveys should seek expert review. A considerable stock of knowledge has been developed on survey construction. For example, survey items are more reliable when they ask behavioral questions and employ behavioral answer categories; when they are asked in a non-threatening way; when answer categories are somewhat subtle ("always, usually, sometimes, never" rather than "yes, no"); for oral telephone surveys, when no more than 4 or 5 categories are used; when questions are short and contain only one idea; and so forth.

Another possible source of reliability problems is faulty recall. This possibility may be a particular concern of agencies planning longer follow-up periods or planning to ask retrospective questions concerning experienced during treatment. However, there appears to be little reason for concern. Duncan et al. (19 81) finds little recall distortion on any but trivial issues with follow-ups of under one year. That distortion is reducedwith repeated interviews as opposed to a single long-term follow-up. Bradburn (1983) recommends the "bounded recall" method to prevent error from telescoping time or forgetting isolated events. The interview begins with clearly anchored facts and works forward or backward out from that point.

One possible reliability problem applies in particular to research on programs for the disadvantaged: some respondents will be learning disabled, have little education, or speak little English and may therefore experience difficulty understanding even well constructed questions. The solution to these problems lies with interviewer training, and with availability of bilingual interviewers where possible.

Of course, as discussed in the main text, sample size is the primary influence on the reliability of particular estimates.

Validity

The area of validity is more problematic than reliability for surveys, although here too some types of early concern have been laid to rest. "Validity" refers to accuracy -- is the conclusion of the research, or the particular measure, true? Are estimates "unbiased," accurate on average? Highly reliable, but invalid, survey questions are capable of consistently generating inaccurate information, like a watch which runs 10 minutes slow and therefore consistently shows the wrong time. Validity problems are especially serious because larger samples cannot overcome them. They simply reinforce our faith in the biased estimates. Major validity problems for surveys such as the gross impact model proposes may be divided into "response effects" and bias from nonresponse.

Response Bias. Bradburn (1983) provides us with an excellent review of the large body of research on response effects. Weiss's (1975) review is also especially useful in the present context since she writes specifically about interviewing for program evaluations. Response effects occur when the respondent decides, consciously or unconsciously, to report incorrect information. Questions can "lead" the respondent, or respondents may feel pressure to answer in the way they think the interviewer wants them to answer. Even without pressure, respondents may give the answers they believe are socially desirable.

However, Bradburn's review of the large number of studies directed toward this problem are reassuring. Although such bias can occur, especially when "loaded" questions are asked, signs of serious response bias are rare, and typically occur only in situations which are obviously problematic. Minor response bias is likely: slight inflation of earnings reports, or a slight movement of satisfaction ratings in a positive direction. But these are unlikely to be serious. One of the most reassuring conclusions comes from Weiss's review of research examining the selection of interviewers. Although a stiff, formal interviewer style tends to generate higher refusal rates and more socially desirable answers, other factors such as age, sex, education level, or identity as indigenous to the area have no such predictable effect.

There is evidence drawn from comparisons of CETA follow-up survey data and Unemployment Insurance records pointing to a mild but rather consistent over-reporting of income, one of the most basic outcomes in this research area. Zornitsky et. al. (1985b), find three differences between survey respondents and UI records in Baltimore:

1. There were some apparently random differences, which may occur because of reporting time period incompatibility or the lower precision of survey income estimates.

2. A considerable group of survey respondents reported earnings where none at all appeared in UI records. This may stem from overreporting (response bias), from work with employers who did not participate in the UI system, or from individuals who had moved out of the state and were therefore not included in the Maryland UI system. These latter two interpretations are strong possibilities where the disagreement between the two estimates is so marked.

3. A considerable number of individuals also showed moderately higher earnings when estimated using survey data than when UI records were used. This may represent a tendency for survey estimates to overlook time lost from unpaid leave and the like, or it may represent response bias.

Finally, Weiss (1975) identifies one source of possible response bias which could face some local data gathering attempts. Program staff used as interviewers often resist the interviews as low priority because they are not service. They may therefore be unwilling to separate interviewing from their roles as service staff. This means they may lead the participant to respond in ways emphasizing their need or emphasizing their gratefulness for the program. Similarly, when data are also used to fulfill performance requirements, yet another source of possible bias arises when staff double as interviewers.

Nonresponse Bias. The second type of bias in surveys is more problematic, primarily because we are so little able to specify the degree or even the direction of its effects. For this reason, the issue received greater attention in the main text and less is said here. However, one issue requires further discussion: testing for nonresponse bias and sample adjustments to compensate for such bias.

Aside from prevention, the most common way of dealing with the possibility of nonresponse bias is to test for differences on known quantities between respondents and nonrespondents, and to make adjustments to the sample if necessary. Many data elements should be available for everyone included in the sample of names to be contacted, regardless of whether the contact was completed. These include MIS data on individual characteristics, the activity in which individuals were enrolled, and termination status. After interviews are completed, those interviewed and those not interviewed may be compared on these data elements. In particular, factors known to be associated with employment status or income are important to test in this fashion.

This source of error should be taken seriously. However, even in the face of relatively serious bias, effects on estimates are relatively small. Duncan et al. (1981) find, for example, that less intense tracking efforts increases non-response rates for some groups, especially minorities, more than others, but that estimates from research with greater and lesser bias are identical. An example will illustrate why.

Gender always predicts income, and most surveys experience a somewhat higher completion rate for women than for men. Let us say that a set of completed interviews includes 50% women, but only 40% of nonrespondents are women. Further, imagine that calculations based on completed interviews show women earning an average of $.70 per each dollar earned by men. Let us say men average $6.00 per hour and women, $4.20 per hour. Thus, we would calculate an overall average of $5.10, based on completed interviews (50% men and 50% women) while our best estimate of the average for nonrepondents would be $5.28. With a 70% completion rate, the adjusted estimate for the entire sample would be $5.15, $.05 more than values estimated on the basis of a sample biased by the overrepresentation of women.

Although this bias is small, we might wish to adjust our findings to eliminate it. One method is to report wage findings separately for men and for women. The other is to weight our sample to make calculations behave as if the sample included the correct proportions of men and women. If that true proportion is 45% women, rather than the 50% women in our sample, the weighting scheme can give each female respondent a weight of .9 and each male respondent a weight of 1.1. This transforms the 50-50 split into the desired 45-55 split without changing the sample size on which statistics are calculated. When average wage is calculated using these weights, this particular bias will be eliminated. The problem is that weighting schemes can become complex and affect variance within the sample in ways the analyst may not notice. Therefore, to avoid making the cure worse than the illness, one should move into weighting schemes with caution.

One additional tool is available to those who conduct an employer survey also. Interviews with employers can be completed in over 90% of cases. That means findings from employer surveys can be used to check for differences between participants who respond and those who do not on a subset of the variables which really count - outcome variables. The procedure is the same as just described, except that the adjustment may be made directly to the outcome, job retention.

For example, imagine that our employer survey shows 70% of participants still with their placement employer at 3 month follow-up, but that our participant interviews show a retention rate of 75%. This is a rather large bias which we would wish to correct. The weighting procedure would be used to decrease employer retention from 75% to 70%, while increasing other employment statuses from 25% to 30% of the sample. These adjustments could be made to already calculated outcome statistics, or for findings such as job qualities, by reweighting the sample. Assigning weights of .93 and 1.2 to retained and non-retained participant interviews, respectively, would accomplish that reweighting. If half of those no longer with their termination employer have found new work by follow-up, this reweighting would correct our estimated rate of employment at termination from 87.5% to 85%.

DIFFERENTIAL IMPACT ANALYSIS AS A STATISTICAL MODEL

The differential impact approach compares different treatments (activities, processes, and services) rather than comparing treated individuals to an untreated comparison group. The goal is reliable estimation of the direction and strength of impact each variant has on participants' post-program labor market experiences. These research questions involve testing causal models.

The minimal essentials for causal analysis are an intellectually plausible causal argument, appropriate research comparisons, and appropriate elimination of, or control for, alternative explanations. In non-experimental research, identifying causal relationships is always problematic, primarily because of potential selection bias. Since individuals are not randomly assigned to treatments, we cannot be certain that differences associated with these treatments do not occur because of other differences among the individuals who entered each treatment. However, the same "quasi-experimental" design which makes net impact studies possible also provides the basis for differential impact analysis (Campbell and Cook, 1979). The groups compared are differentiated by the treatment they receive, each acting as comparison for the others.

The best way to approach the validity of differential impact analysis design is to begin with the design of net impact studies, which is widely accepted. The precise model for net impact analyses appears in various forms depending on the author's focus. However, a general specification may include the following:

$$Y_i = a + b_1(ENV)_i + b_2(CHARS)_i + b_3PROG_i + e_i$$

where Y = the outcome, such as earnings or employment status,

a = the intercept and b1, b2, and b3 are sets of regression slopes,

ENV = a set of local labor market characteristics which affect the outcome,

CHARS = a set of individual characteristics which affect the outcome,

PROG = a dichotomy indicating whether the individual participated in the program, and

e = an error term including various components, specific to individual i, specific to the time period, or random.

This approach is well established (e.g., Bassi, 1983; Bloom, 1984). In the above model, the key test is of PROG, which estimates how much participation affected the outcome. The selection bias problem is that individuals were not randomly assigned to states of PROG. The main approach to reducing this problem is adjusting for individual characteristics (CHARS) and labor market environment characteristics (ENV) which may influence both selection and outcome. A second approach is adjusting the statistical estimation procedures used.

How does the differential impact model compare? It also suffers from selection biases. Some individuals self-select into different activities, and service providers may differentially select individuals into different program activities. The model estimating differential affects differs slightly depending on the level of measurement involved. In the simplest case, comparisons may be made between two program activities such as classroom skills training and on-the-job training, or between two service delivery organizations. In such cases, the model is identical to the one specified above except that the PROG dichotomy changes from "participant/non-participant" to

"CT/OJT," or "Agency A/Agency B." The resulting slope estimates which treatment alternative has the more positive impact, and by what margin. As with net impact estimates, bias is reduced primarily by inclusion of CHARS and ENV, but is never eliminated.

A second level of measurement occurs at the service provider level, but measures some specific policy or practice, such as intake diagnostic procedures or use/nonuse of subcontractors. Here, the PROG term in the above model becomes a set of variables, each taking on values observed across service providers in the sample. This level of measurement specificity increases the demands placed on the sampling design and also increases our ability to avoid bias from non-random selection.

To test such measures, the sample must include multiple service providers. This is the research design key to effective differential impact analysis. Only with multiple service providers can we test the effects of service provider policies. With only two agencies, all variables describing them are confounded. Identical correlations will occur between an outcome and any characteristic which varies between the two agencies. However, as the number of service providers grows, so does the number of agency level variables which may be tested. This adds to our model the ability to test various agency policies against each other to avoid reporting spurious agency effects. It also offers added protection from one source of selection bias: agency selection practices which may now be measured and included in the model.

The third level of measurement specificity adds to the PROG term the specific program experiences of each individual participant. These might be called "standardized individual process measures." For each participant in all service providers in the sample, particular program experiences are measured, making available truly individual level data on treatment variation. The major cost here is data collection effort.

The benefits are two. First, we can estimate the precise impact of treatment experiences rather than average impact between agencies. This means the number of service providers in the sample becomes somewhat less critical than with agency level measures. Second, we have yet another tool with which to guard against selection bias: we can measure the route that specific individual took into training. The most powerful example of this in my interviews involves referral to OJT by the employers who then become the OJT employer/trainer. About one-sixth of all OJT employers stated that they had made a firm hiring decision before referring the individual to CETA to see if s/he was eligible for OJT wage subsidy. This route into OJT represents rather full fledged "creaming." Being able to measure such processes greatly enhances our ability to identify selection effects.

In sum, the differential impact approach offers quite good protection from non-random selection bias, although there is, of course, no substitute for experimental design. The differential impact approach therefore offers valid estimates of the direction and strength of impact produced by a wide range of program policies, typical practices, or individual treatment experiences.

INDIVIDUAL AND ORGANIZATIONAL LEVEL MEASURES

Research which estimates impact of several different individual characteristics (gender, education, employment history, etc.) has become commonplace. But how do we combine individual variables such as gender or age with variables measuring characteristics of treatment or service providers? This question can be addressed more easily after defining several types of measures, their unit of measurement, and their unit of analysis. These are listed in summary form in Exhibit B.2.

EXHIBIT B.2. INDIVIDUAL AND ORGANIZATIONAL LEVEL MEASURES

Type of Measure	Unit of Measurement	Unit of Analysis	Example
Individual characteristic	individual	individual	sex, age
Aggregate organizational characteristics	individual	individual or organization	percent female
Individual program experience	individual	individual	length of program
Organizational characteristic	organization	individual or organization	quality control policy

These types vary by unit of measurement and unit of analysis. The unit of analysis is the level (individual, group, organization) constituting cases on which statistical analysis is performed. The measurement unit refers to the thing actually measured. The first two types listed above are both based on characteristics individuals bring with them to programs. They differ only in that the second type is computed as the aggregation of the first. DOL annual status reports take this latter form.

The third and fourth types are measures of program treatment or administrative characteristics. We tend to think of these as always measured and analyzed at the organizational level, but this need not be the case. Policies and typical practices of organizations are measured at the organizational level, but may be analyzed using either organizations or participants as the unit of analysis. When the unit of analysis is the individual, all individuals enrolled through a particular organization take on that organization's value on the variable and the variable becomes something like: "enrolled through a service provider with X policy." Other program variables are measured directly at the individual level because they vary among individuals enrolled through the same organization. Although this matter has not been thoroughly studied, it would be surprising if variation in treatment experience were not greater among individuals enrolled through the same providers than the average variation between service providers.

Service Providers (Organization) as the Unit of Analysis

One approach to integrating individual and organizational variables is to transform all variables to the organizational level for analysis. Characteristics of individual clients and of individual treatments are aggregated into organization averages or percentages, and these are combined with organization level measures of program policies. Outcomes are also aggregated into rates for each organization, so that all analysis takes place at the organizational level. This approach is necessary if individual measures are not available and is therefore used to analyze summary records. However, it is much less powerful and less accurate than individual level analysis.

Individuals as the Unit of Analysis

In this approach, the program experiences of each individual are treated as characteristics of the individual, whether measured at the individual or at the organizational level. Thus, a particular participant may have characteristics including: female, age 22, high school graduate, participant in classroom training, program length of 4 months, non-stipended, trained in a proprietary school, enrolled in an SDA using integrated intake, placed through a service provider with a job club available, and so forth.

Each of these measures can be tested as a possible determinant of post-program labor market success. Just as we can compare young versus old, we can compare attendance at proprietary schools versus others or experiencing ntegrated versus decentralized intake. Because all these variables characterize individuals, multivariate methods may be efficiently used to test them against each other.

APPENDIX C

Appendix C. Selected Implementation Topics

This guide does not provide detailed implementation instructions. However, some topics concerning implementation problems should be considered to help managers make informed judgements about whether to invest in the kinds of analyses suggested here. Most of these have to do with the practical problems and costs associated with conducting surveys. Readers might consult Weinberg (1983) for an excellent systematic review of concrete steps involved in data collection.

PART I.

TOPICS CONCERNING THE ADMINISTRATION OF FOLLOW-UP SURVEYS

LOCATING PARTICIPANTS FOR INTERVIEW

Although low income individuals are highly mobile, no study which plans carefully in advance should experience difficulty completing 70% or more of interviews with participants. The largest influence on successful completion is collecting accurate locator information during enrollment. No amount of after the fact tracking effort can substitute for careful initial recording of addresses and phone numbers for participants and for locators such as participants' family members or other stable individuals well known to them.

A CETA study which was not planned in advance will illustrate (Simpson, 1982). That system imposed little requirement for agencies to collect accurate address information, and no locator information was available. The result was that 15.5% of the sample records included no address or phone number. Another 23.7% had numbers which were no longer valid, reflecting the fact that address information was seldom updated after eligibility review, which had occurred anywhere from 1 month to 2 years previous to termination. The impact of this issue is illustrated in Exhibit C.1. Completions were very high among sample names with relatively accurate address information and very low where tracking efforts had to begin from zero.

An additional impact of accurate locator information not indicated in Exhibit C.1 is the reduction of survey costs. Over one-fourth of sample names including an apparently valid phone number were contacted on the first call. Extensive tracking efforts were required to contact any of the group with no addresses or phone numbers on file, and the majority of those efforts were futile (Simpson, 1982).

EXHIBIT C.1 INTERVIEW COMPLETION RATE BY STATUS OF CETA RECORDS

	Percent of Sample	Rate of Completion
Apparently Valid Phone No. in Records [1]	60.8%	81.8%
Previously Valid No. In Records [2]	23.7%	47.8%
No Phone No. Available in Records	15.5%	33.2%
Total	100.0%	66.2%

---

[1] Includes numbers never reached for confirmation, and message phones.

[2] Includes disconnected numbers, whether confirmed valid or not, and individuals who had left the state, returned to prison, etc.

The second line of attack to improve completion rate is effective tracking. Articles by Barnes, Homans, and Lewis (in Borus, 1972) thoroughly review tracking techniques. Tracking information provided by others who answer at the original telephone number is the single most effective method of locating individuals who have moved. After that, tracking through relatives (i.e., locator names provided with the sample) is most effective. In addition, letters announcing the research and requesting change of address information may be helpful. Other methods include consulting standard telephone directory assistance, calling others with the same last name (if the initial residence is in a rural or small town area), and checking with welfare or local jail officials.

Exhibit C.2 illustrates the relative success of several tracking methods used by Simpson (1982). In that instance, no relatives or other locators were available. In all, 35.6% of all interviews were completed only after tracking, showing that it does play a considerable role in improving completion.

EXHIBIT C.2 PERCENT OF SUCCESSFUL TRACKING USING SELECTED METHODS

Tracking Procedure	Percent
Information from someone at the original number	45.4%
Mailed back new phone number	21.5%
Directory assistance	17.4%
Information from training school	7.6%
Participant called collect, after a mailed request	6.8%
Additional call to CETA offices	1.3%

One final note on this topic: for JTPA programs which place participants on work release, one element of the program locator material gathered before participation should be agreement from the parole authorities to assist in follow-up. Most work release participants have as their only phone number a prison or parole office. Without prior authorization, tracking through these sources is usually impossible.

STRATEGIES CONCERNING COMPLETION RATES

Although minimum expected completion rates will no doubt be established for any follow-up effort, realistic rates depend on factors outside the hands of the interview team, as well as proper tracking and interviewing. For example, with excellent locator information and a largely non-urban clientele (tracking is easier in less urban areas), 80% completion rates may be possible, and a 70% minimum easy to achieve. In other situations, even 70% may be very difficult. With employers, tracking is seldom at issue, so that achieving an 85% or 90% completion should be possible in most cases. This makes 80% an entirely realistic minimum figure.

One difficulty with any particular minimum completion rate is that it might be achieved with relatively "easy" calls only, thus biasing the sample away from participants who move, who work double shifts, who have an active social life, etc. In fact, it is as important that each participant be tracked fully as that a particular completion rate be achieved. Therefore, managers should negotiate an expected level of commitment per sample name over and above the minimum completion rate expected.

The minimal standard set for measuring post-program performance requirements, six callbacks, represents a reasonable quality level and would serve well for other gross impact analysis as well. Calls should be attempted at a variety of times, to insure that individuals who work evenings, travel each weekend, etc. are reached. Such requirements have the effect of equalizing the chances of each individual in the sample to be called, the key element of a representative sample.

INTRODUCTION AND CONTEXT SETTING IN TELEPHONE SURVEYS

Americans are remarkably well trained to answer the phone when it rings and to interact quite naturally over the telephone with total strangers. Nevertheless, telephone interviews should immediately set the proper context for the interview. This includes having the interviewer identify himself or herself by name, by organization, and by purpose. (E.g.: "Hello. My name is Carl Simpson. I'm calling from Research Inc. We have been asked by the S.E. Corner Private Industry Council to conduct a survey concerning individuals trained or employed with assistance from JTPA, the Job Training Partnership Act.")

All participants and most employers should immediately recognize that this is the follow-up call they were told they might receive. For others, such as participants' spouses, this introduction should set a professional yet warm and relaxed tone. The professional quality of the introduction is particularly important when interviewers call in the evening and ask for participants of the opposite sex. The friendly tone is especially important in cases where participants have outstanding debts or where the family is accustomed to negative experiences with official calls.

Some research organizations recommend that letters be sent prior to initial calls, in order to establish legitimacy. However, this somewhat expensive technique is unlikely to have any effect in cases where the individuals being interviewed are already associated with JTPA (Lewis, 1972). In the case of terminating employers who did not participate in delivering services, no previous knowledge may be assumed. However, in those cases an advance letter may raise questions about the participant who was hired. The ability of an interviewer to answer questions as they arise recommends calling rather than mailing first.

After the interviewer has located the individual to be interviewed, the next critical element of the telephone introduction is to establish the neutrality of the interviewer and the research. If possible, the interviewer should be seen as separate from service providers. This prevents participants from politely suppressing criticism or from exaggerating need in order to ask for further services. And it allows participating employers to be less guarded when explaining non-retentions, reporting service levels, or evaluating JTPA services. (E.g. "These interviews are being conducted independently from S.E. PIC. Answers given by individuals are treated confidentially. Only overall results are reported to S.E. PIC. The goal of our work is not to evaluate individuals, but to assess how well current programs are working and how we can improve them in the future.") In cases where the research is conducted in-house, these assurances are more difficult to make believable, and are in fact more likely to be compromised. However, it is important to protect individuals and establish neutrality as much as possible. Research without such guarantees should not be published for general consumption and should be interpreted internally with the recognition that it may be biased.

A final element of the introduction is critical where the respondent is at all hesitant. That is to establish the importance of each individual's participation in the research and to minimize the perceived cost of participating. (E.g. "You are part of a scientifically selected sample of employers who recently hired someone trained with the assistance of JTPA. That makes it very important that we include your experiences in the research. If you have just a moment, the interview is brief...." (If no: "OK, I can schedule a time to call you back anytime between 9 and 5....")

In the case of employer interviews, the relationship between the individual being interviewed and the JTPA participant in question should be established, to insure reliable information. For example:

"You recently hired Jane Doe. I would like to ask you just a few questions about Jane and the job you hired her for. Were you Jane's direct supervisor?"

IF NOT: "Did you work closely enough with Jane to answer a few questions, or would it be better for me to call back and talk to her immediate supervisor?"

In the case of participant interviews, there is usually no ambiguity about who is being interviewed. However, three potential cases of confusion might be mentioned. First, participants may sometimes be confused about what program(s) they have participated in. Second, more than one individual in a household may have participated in a JTPA program. Third, as part of his or her recent JTPA program, a participant may have participated in more than one classroom course or may have been placed with more than one employer. This is especially likely during periods of economic instability, when employers may have to cut back their OJT positions. In each of these cases, the interviewer needs to reach clarity about what experiences are being discussed - for example, which training satisfaction questions refer to.

WHO SPEAKS FOR THE EMPLOYER?

One potential weakness of employer surveys is the possibility that it will be difficult to determine who, in the organization, should be interviewed. In addition, different individuals may be better able to answer questions about the organization and about particular workers. This problem is minor in most, but not all, cases. In particular:

- For local interviews of participating employers, program officers will already have access to both higher level administrators and direct supervisors, developed during the course of arranging the job placement.

- For state interviews of participating employers, part of the employer sample identification should be the name and phone number of the employer's contact person with JTPA, typically the individual who signed the placement contract.

- For small employers, one individual typically owns the business and also supervises employees.

- The most problematic group are large termination employers who did not participate in delivering services. In these cases, a little extra time and good will are demanded to locate a supervisor or personnel officer who can rate the employee accurately. This problem could be largely eliminated by including in agency termination interviews not only the name of the employer but also the name of the individual who hired the participant and/or the participant's supervisor. This information would then be used to guide follow-up interviewers.

PART II.

ELEMENTS OF COST ACCOUNTING FOR IN-HOUSE FOLLOW-UP SURVEYS

It is more possible in a guide such as this to identify the elements of survey costs than to attach specific figures to each element. Costs such as phone tolls, interviewer wages, and overhead vary considerably with local conditions. The following discussion identifies the major elements of cost associated with follow-up surveys, and offers some examples of cost calculations based on the interview instruments included in Appendix E. Estimates for those instruments are based on a modest number of pre-test interviews, but should be accurate within 10-20%.

There is great divergence among different possible ways of calculating costs of follow-up surveys:

1. Out of pocket marginal costs,

2. True marginal costs, including labor reassignment, and

3. Total accounting costs, including a share of fixed costs.

If data collection is contracted, the out of pocket marginal cost to JTPA is by definition equal to the contractor's total accounting costs perhaps plus some profit margin. If it is performed in-house, the physical plant and phone systems are intact. In addition, some staff time may be reassigned, making costs in-kind rather than monetary.

The prime reason for conducting research in-house is budget limitations, making this the situation where specifics of marginal cost planning are most important. In contract situations the contractor will specify accounting calculations, although, of course, managers should be aware of what elements in addition to expertise and reputation are being purchased via data collection contracts.

EXHIBIT C.3. ELEMENTS OF DATA COLLECTION COSTS PER INTERVIEW

Getting Started:

- Sample Selection
- Attaching MIS data
- Recording Data Linkage Codes
- Printing Charges

Conducting Interviews:

- Interviewer Wages
- Telephone Tolls
- Tracking Costs
- Call Back Costs
- Supervisor Wages
- Data Entry and Cleaning
- Utilities

GETTING STARTED

The set of cost categories listed in Exhibit C.3. are grouped together because they all involve data collection and they are all contingent on the number of interviews conducted and/or the length of each. The first two represent modest costs. Lists of participant or employer names must be selected into the sample and recorded in a fashion making them available for interview. It is desirable for each name to be accompanied by locator information and program termination status, including training area, termination status, and employer name, if applicable. This requires either a small amount of computer time or moderate clerical effort.

The second task listed in Exhibit C.3. must also be accomplished at the point of sample selection (or later, but for all in the sample). MIS data available on each participant, or data characterizing each employer serve two purposes. These data are used during analysis. They also provide a basis for assessing how much non-completion bias occurs during the survey, by comparing MIS characteristics of sample members with whom interviews were completed with those not completed. These data must therefore be included for all in the sample.

A third clerical task necessary at the start is the recording of identifying codes, necessary for organizing and merging data sets. (See Chapter 8.)

These three tasks require a period of training and setting up procedures, after which costs probably run in the vicinity of 5-10 minutes staff time per sample name. Since costs may vary widely, agencies planning projects might run pre-tests in order to estimate the number of sample names which may be processed in one hour of staff time.

Printing charges are more straight forward. The number of locator sheets, identifier code sheets, and interview pages can be determined ahead of time. For telephone surveys, the least expensive available method of printing may be used.

CONDUCTING INTERVIEWS

The first step in planning the cost of interviews is estimating the average length of each interview. At least 15-20 pretests are required because interviewers become more efficient after their first 5-10 interviews and because interviews vary in length depending on whether the participant is employed, has a job history, and the like. The estimates shown in Exhibit C.4. are based on pretests of the participant and termination employer instruments included as Appendix E.

EXHIBIT C.4. LENGTH OF PRETEST INTERVIEWS WITH PARTICIPANT AND TERMINATION EMPLOYER *

	Participant Interviews	Employer Interviews
Average if Employed at Termination	7 minutes	6 minutes
Average if Unemployed at Termination	4 minutes	5 minutes
Overall Average Interview Length	**6 minutes**	**6 minutes**

---

* Based on 21 participant pre-tests and 15 employer pretests, using the participant survey instrument and the employer survey instrument for all employers, in Appendix E.

Using the average figures from Exhibit C. 4., and adding time for callbacks and for tracking, for compeletions and non-completions, reasonable estimates for average **telephone time per completed interview** is approximately 8 minutes for participants and 7 minutes for employers. Cost per interview depends on the number which are local calls, on within-state toll rates, and on the proportion of calls out of state.

Interviewer time includes more than phone time. In addition to break time, paperwork and dialing time between calls, instruments should be quickly reviewed after each call is completed, and tracking efforts can be time consuming. These figures add to an estimated **interviewer/tracking time per completed interview** of about 15 minutes for participant interviews and 12 minutes for employer interviews. Thus, if the same individual(s) do tracking and interviewing, one can estimate about four completed participant interviews per hour and five employer interviews per hour, using the instruments in Appendix E. How this translates into dollars depends on interviewer wage rate and on how many other duties are combined with interviewing time.

Most data analysis costs are fixed, depending little on the number of cases in the sample. However, data entry and data cleaning depend directly on the number of completed interviews. The surveys included in Appendix E should average less than 2 minutes to data enter, with verification doubling that time.

Supervisor costs are difficult to assess for small, in-house efforts. The most feasible arrangement is probably to hire and train interviewers carefully, and to have them work relatively independently. A supervisor can train interviewers, drop by for periodic checks, use an extension to listen to interviews periodically, and check completed interview forms for accuracy, but cannot be present at all times. A goal might be to keep total supervisor costs below 1/4th that of interviewer wages, while insuring the quality of the data collected.

FIXED COSTS

If data are collected by contract, overhead costs may range from 50% to 300% of direct, marginal costs, depending on the contractor and what tasks are included. For in-house data collection, most overhead during data collection (phones, office space, etc.) is in place, making calculation of fixed costs for the project somewhat arbitrary.

In addition to data collection overhead, other fixed costs are involved in conducting research analyses. In particular, surveys must be constructed and pretested, a specific sample design must be developed, and data must be analyzed and findings reported. Since these tasks require professional time, they can become rather expensive. In this area, more than any other, marginal utilities come into play, making state-wide research more cost efficient than local projects (regardless of who actually collects the data). It is our hope that the survey instruments included in Appendix E will also reduce fixed costs by providing some of the professional work involved in survey construction.

Exhibit C.5 summarizes some major elements of fixed costs faced by managers deciding whether to engage in analyses built on follow-up surveys.

EXHIBIT C.5. FIXED COSTS ASSOCIATED WITH FOLLOW-UP SURVEYS

Interview Set Up Costs

- Interview construction and pretesting
- Sample design
- Interviewer training

Overhead Enabling Data Collection

- Phone installation
- Physical plant overhead
- Clerical capacity for call-ins, filing
- Data storage and analysis capacity
- Equipment, supplies

Costs of Developing and Disseminating Results

- Data merging and variable modification
- Data analysis, equipment and personnel time
- Write up - professional time
- Printing
- Dissemination workshops, etc.

Interview Set Up Costs.

Unless an SDA happens to have on staff someone trained in survey construction, or wishes to conduct only the most cursory follow-up, the construction of survey instruments should be made a shared project. If a consultant is paid $3000.00 to construct a survey and sampling design, and to train interviewers, the startup cost approximately equals the entire data collection cost of a typical local project. If pre-existing interviews such as those in this guide are used, startup costs are nearly eliminated. However, care must be taken to insure that the instrument adopted will accomplish managers' analysis goals.

Designing an appropriate sample is very easy in the case of local studies. Interviewer training is important and should be done by a professional. However, it is brief, requiring no more than two hours of interviewer and supervisor time during training plus somewhat more active supervision during the first few days of interviewing activity.

Overhead Enabling Data Collection.

This is an area subject to considerable negotiation. Contractors negotiate during the bidding process. Managers performing in-house data collection efforts may need to negotiate figures on several fronts:

1. What true costs are added to existing costs by the analysis project, if any.
2. What proportion of existing overhead is reasonably attributed to the data collection project as a portion of organizational activities.
3. How do type 2 calculations affect administrative overhead calculations and, where applicable, the appropriate level of state or other special financial support for SDA follow-up analysis.

Costs of Developing and Disseminating Results

The final fixed cost is crucial in two respects:

1. It can be rather expensive, since it involves computer programming and professional time, and
2. All too frequently, dissemination is assumed rather than created, with the result that few learn of the results, making the entire project extremely expensive per user.

The survey instruments included in this guide probably require 2-3 days' programmer time to organize follow-up data sets and to modify variables for initial analysis. In addition, up to 3 or 4 days programmer time may be involved in successfully merging data from various different sources, if a full intensity differential impact analysis is to be undertaken. Actual costs vary greatly depending on available computer technology.

In this area also, this guide attempts to reduce new users' fixed costs by including a table detailing the location of each variable measured in the instruments provided, and also providing data modification directions. (See Appendix E.)

Analysis and write-up are also expensive and usually require professional expertise. This may mean contracting that work or hiring an individual whose major duty is providing such analyses.

PART III.

WHO SHOULD ADMINISTER DATA COLLECTION AND ANALYSIS?

The tradeoffs involved in deciding how to administer gross impact analysis may be examined by identifying the major components of the analysis to be administered and the major options for their administrations.

MAJOR OPTIONS FOR THE ADMINISTRATION OF ANALYSIS PROJECTS

Four major options are formed by the intersection of two choices:

- First, should the analysis be administered locally or at the state level?
- Second, should it be administered in-house by JTPA agencies, or through a contract to a private research firm or university research group?

Each of these approaches has advantages and drawbacks, discussed after the tasks to be administered are identified.

MAJOR TASKS TO BE ADMINISTERED

- First, the specific research design, including sampling and measurement, must be established. It must be standard across all service providers and participants in the sample. Therefore, advance coordination and planning are required.
- Second, the specific members of the sample must be identified. Rules must be standardized across all organizations in the research, but the lists of participant names which constitute the population from which the sample is drawn originate at the local service provider level.
- Third, data must be collected. Measurement of program implementation variants is logically a centralized survey of local administrators. Individual treatment program variants are best measured by direct service staff, but may in some cases be integrated into follow-up surveys. Given use of telephone surveys, follow-up data can be collected at state or local levels.
- Finally, data must be computerized, analysis files constructed, and analysis performed. For descriptive usage, any organizational level capable of performing these tasks may appropriately do so. For differential impact analysis purposes, data from all service providers must be combined. Therefore, even if data are collected locally, they must be centralized for analysis.

If the analysis is to be SDA level only, administration will obviously be at the SDA level. However, the central versus local data collection question remains open if the SDA uses subcontractors. If the state is the instigator, funding agency, and major user of the analysis, planning and analysis will clearly be administered at the state level, and data collection is also likely to be controlled at the state level.

Beyond that, the first and last tasks, set-up coordination and analysis, as well as program implementation measures are best pursued by the highest organizational level involved in the analysis, state, consortium of SDAs, or SDA. If that organizational level lacks regular staff, as in a temporary consortium of SDAs, these tasks should be subcontracted. Otherwise, choices remain available.

Also, sample identification and individual treatment variants will normally be collected locally, under the administrative supervision of whatever level is chosen to collect other data. However, some treatment variants could be reported during follow-up interviews.

The list of decisions remaining are whether to collect data locally or at the state level, and whether to subcontract data collection, and perhaps other tasks, or to administer them in-house. Exhibit C.6 suggests how these options fare on four criteria which might be employed to guide these decisions.

SOME CRITERIA FOR DECIDING HOW TO ADMINISTER ANALYSIS PROJECTS

The following four criteria could be used to guide the decisions of whether to collect data locally or statewide, and in-house or via contract:

- What level of **specialized technical competence** is required, and what level is available from each option?
- How **legitimate** the research effort will be. That is, how much credibility will consumers such as political leaders place in findings.
- How much is each approach likely to **cost**?
- How much is **involvement by users** of the analysis results encouraged or discouraged by each approach?

EXHIBIT C. 6. COMPARISON OF DATA COLLECTION ALTERNATIVES

	State Level		Local Level	
criterion:	in-house	contracted	in-house	contracted
Specialized Competence	med-high	high	low	med-high
Perceived Legitimacy	medium	high	low	med-high
Cost	low	high	med-low	med-high
Involvement by users	med-high	Low	med-high	med-low

The ratings in Exhibit C. 6 are generalized and for illustrative purposes only. Even so, they identify some broad trade-offs involved in such decisions. Regarding the subcontractor versus in-house administration option, the major trade-off is between legitimacy and cost. Both actual and perceived technical competence will (typically) be greatest when data collection is subcontracted. Various conflict of interest possibilities are prevented, and staff qualifications become part of contractual obligations. On the other hand, active involvement on the part of potential users is typically reduced through subcontracting, which means the money is more likely to be spent and forgotten.

The major drawback to the subcontract approach is cost, which is usually higher when research is contracted to organizations which must cover a quite high fixed overhead.

Research firms do have the advantage in two cost areas. First is the ease of flexibly assigning fully qualified personnel part time to small projects. The second is that in-house efforts may suffer high nuisance costs. If the research effort is a one-time only add-on for those administering it, much of the administration process will be new and outside the training of those in responsibility. In such cases, felt costs can become high and the reliability of the research can suffer.

Nevertheless, the major trade off involved in the decision of in-house versus subcontracted data collection is cost versus expertise and legitimacy, with the balance between these much affected by state or local staff qualifications. If the analysis is to be used primarily within the JTPA system, the less expensive in-house option is probably best. If, however, the analysis is also pointed to a wider public, the legitimacy issue may be a primary one. Obviously, if insufficient in-house expertise exists to mount a project at all, a contract approach is called for.

In the decision between state and local data collection, the comparisons in Exhibit C.6 lean toward the state, with local administration likely to lack in-house expertise and legitimacy.

In the case of in-house data collection, considerable economy of scale accrues to state level administration. The state of Massachusetts confirms this expectation, by reporting a steady progress toward centralized planning, collection, and analysis of follow-up data over several years of follow-up data collection. It appears that if the research is to be statewide in scope, nearly all factors press toward state level administration of the entire project.

However, the apparent inevitability of state level administration, and the concommitant sense that local data collection is unwise or inefficient requires that a "pitch" be made for the feasibility of local administration of gross impact research. The relative advantages of state level data collection rest on the assumption that a full scale, standard priced analysis is being funded. However, if one assumes an interest in doing some analysis "on a shoestring," or if only one SDA wishes to pursue an analysis, the obstacles are not as formidable as might be thought. In particular:

- Coordination costs are reduced locally. and the SDA has direct control over subcontractors, helping to insure their cooperation.
- One major fixed cost which normally recommends economy of scale is the cost of development - sample plans, survey instruments, etc. One purpose of this guide is to offset as much of that fixed cost as possible.
- Local data collection costs can be the lowest of all, given some administrative commitment to flexibility.
    - -- Most local phone calls will be toll free.
    - -- If work load allows and flexibility for comptime exists, regular staff may subsidize the data collection effort.
    - -- Supervision and training time are marginally higher locally, but these may not have to be real dollars if staff flexibility is great enough.
    - -- Assuming a locally computerized MIS, data entry can occur locally as readily as elsewhere.
- One intangible cost of local in-house data collection is suspicion concerning its legitimacy. This may matter little if the analysis is purely for internal uses. If it does matter, this cost can be minimized by explicating all design decisions in writing, by hiring and training part- time interviewers rather than using regular staff, and by retaining a credentialled consultant to review operations periodically.

- The matter of data analysis is somewhat more likely to be problematic for local efforts. Local computer software and storage flexibility may limit ability to merge data from different sources and to perform any but very basic or pre-set calculations. If analysis goals are descriptive only, problems may be minimal. However, differential impact analysis may be impossible.

APPENDIX D

APPENDIX D. ILLUSTRATIVE APPLICATIONS

Before deciding whether to commit resources to a gross impact analysis project, it may be helpful to see concrete examples of the types of questions which can be asked and of concrete findings. The examples below are taken from one study of CETA OJT (Simpson, 1984a). They are not intended to be exhaustive or even to represent the major applications of the model. They simply offer brief, concrete examples.

EXAMPLES OF DESCRIPTIVE APPLICATIONS

OJT's Responsiveness to Local Labor Markets

Since OJT placements cannot be arranged unless the employer has a job opening, most OJT placements will occur in jobs or with employers who are stable or growing (Manpower Training Institute, 1976). This claim for OJT is important, since other JTPA services not involving placement with an employer can easily err by training in low demand areas.

Descriptive findings from one question asked of OJT employers are enough to increase our confidence that this logical assertion holds much truth, although these findings cannot be definitive. The question asked: "Has the volume of your business been growing or declining during the past 3 years?" The response categories for this question, along with the number of employers answering in each category are shown below.

EXHIBIT D.1. GROWTH RATES OF OJT EMPLOYERS DURING THE PREVIOUS THREE YEARS

Grew a great deal	33.7%	80.2%
Grew a little	27.8%	
Stable	18.7%	
Declined a little	8.8%	
Declined a lot	11.0%	

TOTAL NUMBER INTERVIEWED = 517

NOT INTERVIEWED BECAUSE NO LONGER IN BUSINESS = 48

Bearing in mind that this was a period of recession, when the great majority of businesses were cutting back or barely holding on, these findings suggest that OJT is quite good at locating employers with whom participants have at least a chance of remaining employed. On the other hand, it also appears that OJT assignments were made to some businesses that were in serious trouble and did not survive.

It is the weakness of descriptive research that we cannot offer more precise conclusions than this. However, compared to designs that can do so, this finding was generated at very modest cost.

How Long Do OJT Participants Retain Their OJT Jobs?

One advantage of OJT is that the employment is found as part of the training; one need only hang on to it. This produces two separable questions: how many hang on during the period when the OJT contract provides employers with reimbursement, and how many hang on after that incentive ends? The graphs on the next page answer these questions descriptively, for one particular sample.

Over four-fifths (83%) of OJT assignments continued throughout the contract period. However, 30% of participants left or lost those jobs within the first month following the end of their contracts. From that point on, OJT retention declined at a very slow rate, with a few jobs lost or left each month, and most of those replaced with other, post-OJT positions.

Again, these descriptive findings do not offer us interpretations as to why the only period of rapid decline is during the month following the contract end. However, they clearly illustrate for OJT managers that sheer termination figures are not adequate to indicate how successfully the OJT program is serving either participants or employers.

Exhibit D.2

Percent of Participants Retained with OJT Employers

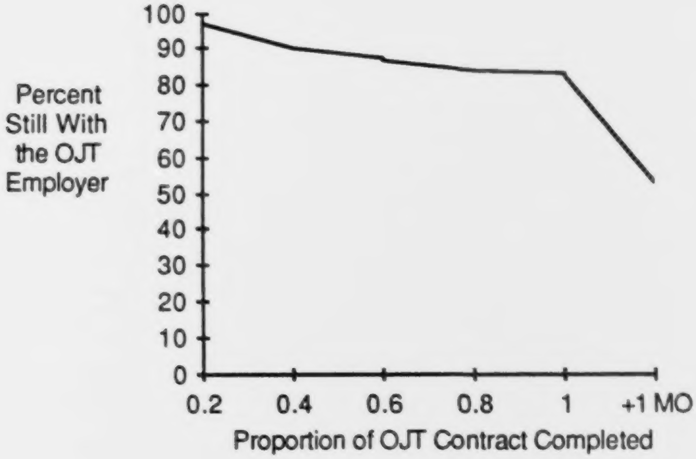
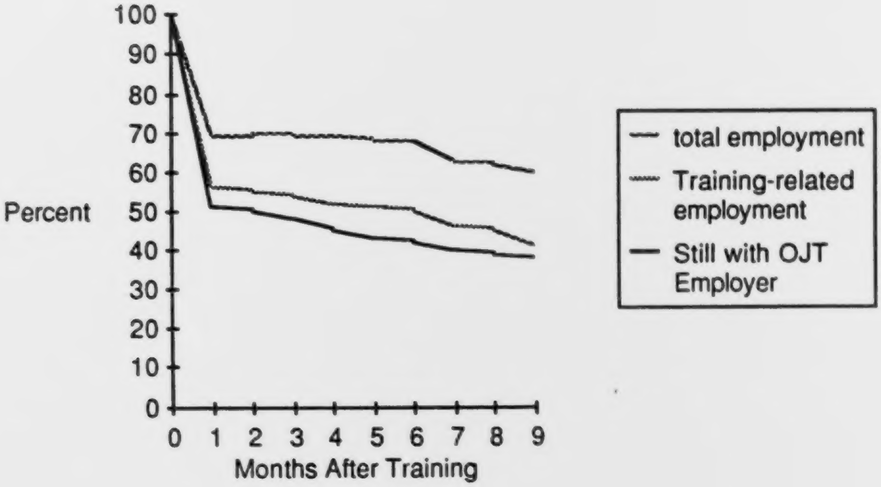


Exhibit D.3

Post-Program Employment Status, OJT Participants



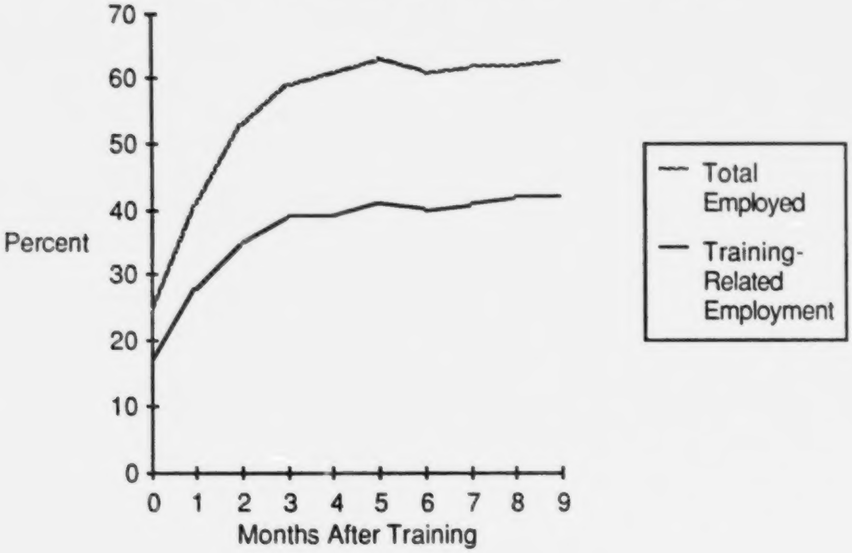
194

What, By Contrast, Is the Post-Program Employment Pattern For CT?

The value of extending descriptive outcome data into the post-program period is illustrated by comparing the OJT figures above with similar findings for classroom training. These findings were not collected during the same year, and have not been adjusted for economic changes or differences in participant characteristics. They may not therefore be compared explicitly. Nevertheless, the shape of post-program employment patterns is so clear in these two cases that our understanding of employment patterns during the immediate post-program period is enhanced by these unadjusted observations.

Exhibit D.4

Post-Program Employment Status, CT Participants



What Evidence Do We See of Skill Utilization From OJT Instruction?

The degree to which OJT produces transferable skill training is an open question (Taggart, 1981). One approach to the question of skill transfer is to ask what proportion of post-training positions remain in the training field, and in particular, what proportion of jobs taken by those no longer in their OJT positions are training related.

EXHIBIT D.5. TRAINING-RELATED AND OTHER JOB STATUSES SIX MONTH AFTER OJT

Employment Status	Percent of all	Percent of Employed	Percent of Non-OJT Employed
OJT Job	39.9	60.3	—
Other, Related	8.3	12.5	31.6
Other, Not Rel.	18.0	27.2	68.4
Not Employed	33.8	—	—

Exhibit D.5 illustrates both the value and the limitation of descriptive analysis. On the positive side, simply by having post-program figures, and by measuring more than income levels, we gain information about how OJT works. As long as OJT retention remains high, OJT produces considerable skill utilization. However, when we focus on skill transfer by looking at skill usage among those no longer with their OJT employers, the training utilization ratio falls from nearly 3 out of 4 jobs to less than 1 out of 3. On the negative side, this finding leaves us with the need for more information. How does this compare with other programs, controlling for participant characteristics and economic conditions? These questions demand differential impact analysis.

How Satisfied Are Employers with Program Services and Participants?

Satisfaction may be measured by direct questions for which descriptive answers are the appropriate finding to report. The particular measures reported here used a 0-10 scale, with 0 representing the "lowest possible" and 10 the "highest possible" evaluation.

One may examine overall levels of evaluation. Comparing the same satisfaction measure for different parts of the OJT program offers additional insight. For illustrative purposes, three employer evaluations are summarized in Exhibit D.6. These are reported for only a subset of 331 employers who were new to the OJT system at the time they hired the participants about whom we interviewed them. Repeat employers were eliminated from this analysis because they were by definition satisfied enough to use the program again. Other employers in their cohorts who were less satisfied might have selected not to participate again. Therefore, new employers give the most accurate picture of how the OJT program was affecting employers. Posing questions carefully in this manner strengthens descriptive findings.

EXHIBIT D.6. EMPLOYER SATISFACTION WITH SELECTED ASPECTS OF OJT PROGRAMS

	The Target Being Evaluated		
Evaluation Score	OJT Representative	Client Screening	Overall CETA OJT
9-10 (High)	56.5%	11.8%	27.4%
7-8 (Above Average)	30.6	33.1	41.1
5-6 (Average)	8.5	32.6	17.7
3-4 (Below Average)	2.5	11.8	8.8
1-2 (Low)	2.0	10.7	5.1

These findings suggest a picture of excellent staff relations with employers pulling overall employer satisfaction up, while less satisfying client screening depresses it. We do not know that these differences represent effects on overall satisfaction, but the possibility is logical. One other guess is that OJT contracts typically do not occur until employers develop positive evaluations of a representative with whom they will work. That guess is mentioned in order to emphasize that descriptive data do not test interpretations, yet they can alert us to important questions calling for interpretation.

Under What Conditions Are OJT Jobs Ending?

The reasons why OJT jobs end tell us something about the fit between participant and employer needs. If most were to end because participants were fired, it would appear the jobs were too demanding or the employers insufficiently reimbursed to afford the inconvenience of poorer than average workers. If most end from quits, either the jobs are low quality or participant employment blocks in the area of personal stability have not been dealt with. Therefore, simple descriptions of the reasons given by employers for firings or by participants for quitting, can offer some interpretive direction. Examples appear in Exhibits D.7 to D.9, below.

It may be surprising that more participants quit than were fired, even during a rather severe recession. The reasons given by participants and employers for their decisions offer rich information about what patterns occur in participants' work lives, but no information whatever about how the OJT program impacts them, compared to what their work lives would have been like without the program or if they had enrolled in a different program.

EXHIBIT D.7. RELATIONSHIP TO OJT JOB AT FOLLOW-UP

Quit	31.0%	
Fired	21.7	
Laid Off	16.1	
Still In OJT Job	31.1	
	100.0%	N=881

EXHIBIT D.8. AMONG THOSE FIRED, REASONS GIVEN BY EMPLOYERS

Unable to do the work	31.1%	
Attitude/personality problem	29.9	
Unreliable	14.7	
Personal problems hurt work	7.9	
Personal habits (e.g., drinking)	6.7	
Theft	6.1	
Careless with safety	3.6	
	100.0%	N=190

EXHIBIT D.9. AMONG THOSE QUITTING, MAIN REASON GIVEN BY EACH PARTICIPANT

Not get along w/ boss	21.5	
Better job offer	19.3	
Felt got no training	8.8	
Family change	7.0	
Bored with job	5.4	
Behind in work	4.8	
Not get along w/ workers	3.9	
	100.0%	N=271

EXAMPLES OF DIFFERENTIAL IMPACT ANALYSIS APPLICATIONS

We turn now to examples of differential impact analysis. These often begin with description and move on to multivariate tests. Although full tables of multiple regression results are not presented, each illustration is the result of multivariate analysis procedures unless explicitly noted.

The Value of Measuring Multiple Outcome Variables

For any given differential impact test, one outcome variable must be selected. That is the nature of multivariate analysis. Before selecting one outcome for use in these examples, the value of having several available may be illustrated.

Taking for these examples only the simplest of outcomes, employment status, we may examine several different dichotomies, defined in Exhibit D.10. Our analysis measured a large number of individual and programatic variables, including data from employers. When these are jointly tested, we learn how much combined power they have to explain each outcome. That gives an indication of which outcomes programs are most able to affect, and therefore, which outcome is most able to help us detect effects of treatment variants. Selecting the outcome most sensitive to program variants enhances the value of findings to program managers.

EXHIBIT D.10. PROPORTION OF VARIANCE EXPLAINED IN SEVERAL OUTCOMES.

Outcomes Including All OJT Participants:

All Employment Versus Not Employed	.32
OJT Retention Versus All other Statuses	.42

Outcomes Including Only Participants No Longer With Their OJT Employers:

Post-OJT Re-employment, All Fields, Versus Not Employed	.14
Post-OJT Re-employment, Training-Related Versus Not Related	.21

The key finding in Exhibit D.10 is that we have much greater ability to explain whether one keeps the OJT job than whether one gets other work. Our ability to explain total employment lies between these others because it is composed from them. This finding occurs primarily because:

- The program variants included in the analysis (but not the control variables) have greater impact on retaining the OJT job than on finding other employment, and
- Employer ratings of the participant have great power to predict whether the participant remains with that employer, but almost no power to predict subsequent employment.

It is appropriate to make these statements explaining these findings because the differential impact analysis that produced them tested a wide range of program and control variables.

Interpretations are bounded by the particular range of measures available and by any number of unmeasured factors at work in any one study.

Does the Route Through Which OJT Positions Were Developed Affect OJT Retention?

Intake procedures and "creaming" are for for the most part very difficult for research to detect. One approach, however, is remarkably simple. There are several routes into an OJT position which can be identified with information from OJT participants and employers. Each involves a different degree of OJT program control over OJT placements and, conversely, different degrees of employer selection. For the most part, employers presumably attempt to select the cream.

After identifying routes to measure, the most basic findings for this analysis are:

- The number of OJTs developed through each route.
- Descriptive OJT retention rates for participants taking each route.
- Multivariate estimates of the impact of taking each route. Multivariate tests were performed by constructing a "dummy variable" for each route.

EXHIBIT D.11. PROFILES OF OJT CONTRACTS DEVELOPED THROUGH DIFFERENT ROUTES *

Route	Percent of OJT Participants	Percent OJT Retention	Multivariate Effect?
A. Agency referral to employer who took only OJT applicants	18.3%	39.8%	no
B. Agency referral to employer who interviewed OJT and others	18.8%	34.4%	no
C. Self-directed job search by participant after OJT eligibility	17.1%	26.8%	no
D. Employer initiated, eligibility determined before firm hire	28.5%	31.2%	no
E. Employer initiated, firm hire before eligibility determined	17.3%	43.2%	17.3% higher

---

* Figures at follow-up interview, 6-10 months after OJT contract termination.

The difference shown in Exhibit D.11 between descriptive OJT retention figures and multivariate estimates of the unique impact of each different route into OJT placements provides an opportunity to discuss how such differences arise and what implications they hold for program managers who view these results. Descriptively, the two most successful options are the two where service providers exercised most and least control over the placement (options A and E). In the latter case, the descriptive difference carries over to a unique multivariate effect. That means no other variables we were able to measure could explain away this difference. Therefore, a differential impact analysis (or net impact analysis) which did not include such a measure would produce upwardly biased estimates of OJT program impact.

Correlates of taking route E show these individuals having greater pre-program employment experience, feeling less need to take just any job for income (they also had not initiated contact with CETA), and being enthusiastic workers. These correlates are indicators of the selection advantage gained by employers who utilized this option. This finding also has implications for service quality level, since in these cases creaming is extreme. The entire hiring process was complete before eligibility was determined, so that in many of these cases, the OJT reimbursement was a windfall, not an incentive. Because this effect holds up in multivariate analysis, we have more faith that the route into OJT, not other participant or program characteristics, make the difference.

In the case of routes A through D, the descriptive differences recorded above are "explained" by other variables. The differences still exist, but the causes of OJT retention are better expressed in terms of other variables which overlap with route into the OJT.

This discussion illustrates several qualities of differential impact analysis.

- First, when questions are asked properly, even initial descriptive results can stimulate analysis, in this case, concerning issues of creaming, level of service, and level of agency control over the OJT development.

- Second, multivariate analysis does tend to change the focus off interpretation. The same descriptive findings are being explained in terms of the variables most directly responsible for the observed differences. For example, if agencies wish to continue using self-directed job search by clients, they may improve OJT retention by performing a quality review over each placement before accepting it.

- Third, multivariate analysis is considerably more complex both to conduct and to report.

- Finally, the more factors one has measured, the more possible explanations may be tested, to identify more fundamental causes underlying interesting descriptive findings.

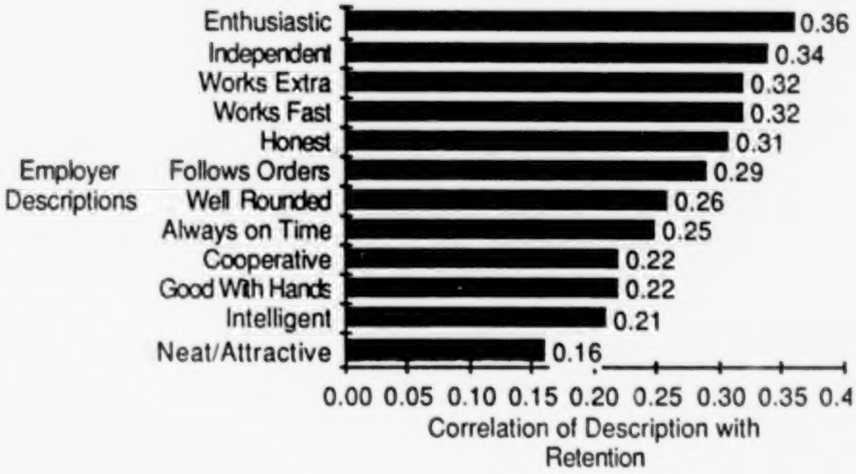
Example of Findings Useful for Marketing and for Planning: When is Cream not Cream?

We asked employers to rate their OJT participants "as workers" on several dimensions. Having these descriptions, and knowing which participants were retained, we can analyze which ratings have greatest power to explain who retained the OJT job and who did not. Thus, we can develop an empirical profile of the "cream" everyone seems to want, not defined in terms of who is employable, but in terms of who retains OJT jobs once the agency intervenes to help make them employable through hiring incentives. Figure D.12 illustrates how close the connection is between each rating and likelihood of OJT retention: the longer the bar, the higher the association.

The four strongest associations shown in Figure D.12 say a great deal. Employees last better when they do lots of work (fast and extra, when asked), and when they do it with minimal burden on supervision (independently and with enthusiasm). Some other factors often mentioned by vocational educators (being on time, following orders, being intelligent) are less closely associated with retention. When all these employer rankings are tested in a single equation, we can identify a smaller set which more parsimoniously summarize the effects of the entire set. In this case, two variables, working enthusiastically and working fast, provide the minimal definition of "cream," at least as regards retention for this group.

Exhibit D.12

Correlations Between Employer Descriptions and OJT Retention



These findings carry implications regarding agency pre-referral diagnosis of participants and regarding which factors are most important to observe if employers are visited during OJT contracts. In addition, they combine with other differential impact findings to emphasize the value of diagnosing OJT assignments which are congruent with both employer and participant needs, rather than "creaming" in terms of general categories of employability. These other findings, not reported in detail here, make it clear that one of the quickest ways to undermine the worker enthusiasm employers want is to place participants in jobs of lower quality and complexity than they are capable of. Thus, attempting to "cream" for the employer, by always referring the most qualified person, may backfire. Workers needs to be qualified enough to learn how to do the work quickly. However, if they are over-qualified, they may lose enthusiasm for the job.

An Analysis of Service provider Policy

Thirty-two service providing organizations were measured as part of our Washington OJT research, with their responses attached to the records of participants who passed through each provider. To illustrate the analysis of implementation variants, findings for one of the many agency policies measures are displayed below in Exhibit D.13.

Agency directors reported their policies regarding the minimum wage for which they would write OJT contracts. Agencies imposing a minimum OJT wage of $4.00 had considerably higher retention rates than others, and those setting a $3.50 minimum had slightly higher retention rates than those accepting minimum wage positions. The first difference carries over into multivariate differential impact findings; the second does not.

EXHIBIT D.13. POST-PROGRAM EMPLOYMENT RATES BY POLICY REGARDING MINIMUM WAGE FOR WHICH OJT CONTRACTS WOULD BE WRITTEN

Minimum Wage Allowed:	OJT Retained	Other Job	No Work	Number Of Cases
$3.35/prevailing	26.8%	36.4%	36.9%	493
$3.50	31.9	26.4	41.8	91
$4.00	39.2	34.9	25.9	189

After controlling other factors in multivariate analysis, service providers with policies demanding a $4.00/hour minimum have higher OJT retention rates than others. However, the analysis also reveals that much of this descriptive difference is accounted for by related factors, a finding which redirects our interpretation. These factors are: a) how complex the job is, with more complex jobs more often retained, b) the length of training offered, with longer training producing higher retention, and c) employer ratings of participants as cooperative and enthusiastic workers, with more enthusiastic workers more often retained.

A Concluding Note

Not reported anywhere in chapter 4 are the many measures in the Washington OJT study which proved to offer no useful information. In the world of research with human beings and human organizations, the inability to explain much is the norm. This is true with differential impact analysis as with other research. That is, in fact, as it should be. The day research "explained all" would be the day before totalitarian rule began. Nevertheless, as illustrated in this chapter, differential impact analysis offers opportunity for managers to receive some useful and reliable information on the probable impact they could have on JTPA outcomes by developing program services in particular directions.

APPENDIX E

APPENDIX E. MEASUREMENT INSTRUMENTS

The best state or local analyses will tailor measures to the needs and interests of the program managers and others who will use the results. However, in view of the considerable start-up costs of tailoring such instruments, this guide includes a set of "off-the-rack" measurement instruments which may be used with minimal modification. They include the most basic individual treatment measures discussed above. Insofar as possible, they are "ready to go."

CRITERIA UNDER WHICH THESE INSTRUMENTS WERE DEVELOPED

The participant survey included in this appendix begins with the three questions which will probably be required post-program performance measures by the time this guide is published. Barring unforseen changes in the form of these requirements, the participant survey should be completely compatible with DOL requirements, allowing the two to be combined when analysts wish to do so.

The data collection instruments which follow are intended for adoption by organizations with little or no experience with survey research. They are constructed to require relatively little training, and to generate reliable answers even when conducted by non-experts. In addition to criteria which apply to the construction of any survey research instrument, the factors listed below were given particular attention, in the hope of making these instruments "user friendly." They are intended to be easy to read, conversational (to set easy rapport), complete, unambiguous, easy to data-enter with minimal error, and also to serve as a ready made codebook for data analysis.

All questions are completely written out, so that they may be read verbatim. Further, they are written in a conversational tone, so that the interview can establish rapport even where the interviewer feels some discomfort.

With very few exceptions, all answer categories are specified, and most are read aloud. When the entire range of possible answers is read aloud, the respondent feels no ambiguity about the type of answer which is called for, and no interpretation is required by an inexperienced interviewer.

Where nearly all responses will be easily recognizable, survey time is conserved by avoiding the process of reading answer categories aloud. However, this occurs only where doing so is unlikely to introduce error. An "other" category is included for ambiguous answers.

Questions flow into each other in a natural, conversational way, and the interviewer is given explicit instructions regarding when to skip questions, and how many to skip. Where possible, for short skips, arrows direct the interviewer visually. Where the skip involves turning one or more pages, instructions clearly indicate the next question to be asked.

Two categories of text in the survey are clearly distinguished: words to be read aloud and instructions to the interviewer or answer categories not to be read aloud. All lower case text, whether part of the question or answer, are to be read aloud. No upper case words are to be read aloud.

The survey is also constructed so as to facilitate quick data entry and to reduce data entry error:

- Where possible, answer categories are indicated by numbers printed in the survey. These numbers are circled, clearly marking the data to be entered.

- All data are computer entered directly from the survey forms. Column locations are shown at the right margin, allowing data entry staff to work efficiently. In addition, variable names are also suggested in the right margin of each survey.

- Questions flow vertically, preventing confusion during data entry.

The survey also includes a set of variable names at the right margin. The presence of these names on the survey itself and immediately adjacent to computer columns prevents errors during the construction of the data set. More importantly, it also transforms the survey instrument into a codebook, a list of variable names and the measures associated with them which helps the analyst organize the analysis.

These surveys are laid out with the goal of conserving printing costs, but only when doing so is unlikely to reduce measurement reliability and is unlikely to increase data entry error caused by crowding items too close together. Survey instruments used by professional organizations sometimes appear formidable. This need not be the case, however, and a prime goal of these surveys is that they be user friendly.

Questions which are not asked, or which are not answered are simply left blank. In cases where some respondents may not know how to answer the question, a code of 9 is uniformly used, next to the shorthand DK, for "Don't Know."

Where the instruments are not clear in and of themselves (e.g., where variables must be coded, interpreted and categorized, after the data are collected) instructions are included at the outset of Chapter 11.

Finally, although far from least importantly, the participant survey included here is compatible with the DOL post-program performance requirements which were being proposed at the time this guide was written.

SPECIFIC INSTRUCTIONS FOR READING FOLLOW-UP INTERVIEWS

Some examples of questions from the follow-up survey included in this chapter indicate how easily the survey format can be followed. Question 2, below, might produce the telephone dialogue immediately following it. (**All lower case text, whether part of the question or answer, are to be read aloud. No upper case words are to be read aloud.)**

2. What is the main reason that job ended? (CODE INTO CLOSEST CATEGORY)

FIRED	1	QUIT, DID NOT LIKE WORK	5
LAID OFF, CUTBACKS	2	QUIT FOR BETTER JOB	6
END OF TEMPORARY JOB	3	OTHER	9:
QUIT, HAD PROBLEMS	4		

INTERVIEWER: "What is the main reason that job ended?"

PARTICIPANT: "Well, I hurt my back and had trouble with the medical bills, and I finally had to quit and move back to my parents' place."

INTERVIEWER CIRCLES THE NUMBER 4 AND READS THE NEXT QUESTION.

Another question, question 13 in the survey, is read in its entirety:

13. Assuming you wanted to, how likely is it that you could keep this job for the next year? Would you say it is . . .

Very . 1 somewhat . 2 a little bit . 3 or not at all likely . 4

INTERVIEWER: "Assuming you wanted to, how likely is it that you could keep this job for the next year? Would you say it is: (PAUSE) Very, (PAUSE) somewhat, (PAUSE) a little bit (PAUSE) or not at all likely?

PARTICIPANT: "What was that in the middle... after very?"

INTERVIEWER: "Somewhat"

PARTICIPANT: "Yeh. Somewhat"

INTERVIEWER CIRCLES 2 AND ASKS THE NEXT QUESTION.

Another question illustrates a space saving technique which breaks the rule that all questions are fully spelled out, but should cause no difficulty.

11. Does this job offer	NO	YES	DK
a. Health insurance?	1	2	3
b. A retirement program?	1	2	3
c. Paid vacations?	1	2	3

INTERVIEWER: "Does this job offer health insurance?"

PARTICIPANT: "No. I sure wish it did."

INTERVIEWER CIRCLES 1 AND ASKS: "How about a retirement program?

PARTICIPANT: "No, only Social Security!

INTERVIEWER CIRCLES 1 AND ASKS: "Paid vacations?"

PARTICIPANT: "Yes. One week to start and two next year.

INTERVIEWER CIRCLES 2 AND ASKS NEXT QUESTION.

One question initially appears somewhat complex to record, but represents a good way to keep the interview conversational without reducing reliability. Different jobs pay in different time units, making it awkward (and somewhat unreliable) to require any one unit when asking about pay. The following example shows the data recorded (in handwriting) for one individual who answered "$5.50 an hour."10. How much does this job pay before deductions?

AMOUNT $______.__ DK 9999999

PER: HOUR . 1 WEEK . 2 MONTH . 3 YEAR . . 4

OTHER . 5 ____________________ DK 9

In the example above, computer instructions can transform the data into compatible forms more reliably than the interviewer or participant could. Instructions for doing so are included in Chapter 11.

SOME READY MADE DATA COLLECTION INSTRUMENTS

The final elements of this guide are several data collection forms. The first is a ready-made participant follow-up survey. Users can easily add or delete questions, as long as the changes are made without disrupting the organization of skips within the survey. If no changes are desired, this survey could be xeroxed and used as it stands.

In addition to the survey itself, a form for recording identifiers is included. Specific coding systems must follow local conventions. The recording form inlcuded here defines computer fields which are long enough to accomodate nearly any coding scheme.

The participant form is followed by a set of three employer interview forms. The first of these measures employer outcomes for all employers. The section suggests measures characterizing employers themselves. These are appropriate for all employers. The third measures a wider range of possible costs and benefits applying only to participating employers.

Each of these ready-made forms -- the identifier form, participant survey, and employer surveys -- specifies data entry to a separate record. The entire set can therefore be used without difficulty merging data,and portions of the set can be used without confusing data entry.

In addition to ready-made employer and participant surveys, an outline for agency-based individual treatment records is offered here. It includes some fully specified questions, but others which require completion by those conducting the research. Since it is intended primarily for the recording of facts concerning the services each participant receives, there is little necessity for questions to be asked in a particular way. Often, no question will be asked at all. Instead, program officers will record information on services as they occur. Consistent with the conventions established here for reading surveys, most questions on this form are printed using upper case only.

This form is intended as a stimulus and framework rather than as a completed product. Therefore, no data entry columns or variable names are included. It is recommended that those who adopt other instruments from this guide follow the variable naming convention of naming measures on the completed form as a separate series, perhaps beginning with the letters AT, indicating Agency Treatment variables. Similarly, data should be entered on a separate line (card image) beginning with the participant I.D. number and the card number.

This instrument may also serve as stimulus to those planning to include questions concerning individual treatment variants as part of an extended participant follow-up survey. Since relatively few questions can be asked using that approach, it is likely that many which states or SDAs would wish to ask are included here, although in a different form.

One recommended element of the agency individual treatment record is not offered in the instrument below: characteristics applying to particular program activities. These are valuable to study, and are discussed in several locations throughout this guide. However, they are beyond the scope of the limited set of measurement instruments included here.

PARTICIPANT 13 WEEK FOLLOW-UP SURVEY INSTRUMENT

MASTER IDENTIFIER COVERSHEET

NAMES (ADDRESSES, PHONES, LOCATING INFORMATION ON SEPARATE SHEET)

PARTICIPANT NAME: ______________________________

MAJOR TRAINER/EMPLOYER NAME: ______________________________

TERMINATION EMPLOYER: ______________________________

			CARD 1	DATA ENTRY COLUMNS:
RECORD AS INFO. BECOMES AVAILABLE	PARTICIPANT I.D.	______________	PARTID CARD "1"	1-9 10
	SDA I.D.	______	SDA	11-12
	SUBCONTRACTOR I.D.	______	SUB	13-14
	PROGRAM OFFICER I.D.	______	OFFICER	15-16
	CLASSROOM TRAINER I.D.	______	CTID	17-19
	PARTICIPATING EMPLOYER I.D.	______________	EMPID	20-23
	TRAINING FIELD	______	FIELD	24-26
	ELIGIBILITY DATE	YEAR: 198__ WEEK (1-52): ___ ___	YEAR1 WEEK1	27 28-29
	ENROLLMENT DATE	YEAR: 198__ WEEK (1-52): ___ ___	YEAR2 WEEK2	30 31-32
	TERMINATION DATE	YEAR: 198__ WEEK (1-52): ___ ___	YEAR3 WEEK3	33 34-35
	FOLLOW-UP INTERVIEW DATE	YEAR: 198__ WEEK (1-52): ___ ___	YEAR4 WEEK4	36 37-38
	INTERVIEWER I.D.	______	INTID	39-40

DATA ENTRY

JTPA 13 MONTH FOLLOW-UP PARTICIPANT INTERVIEW FORM

FOR REFERENCE DURING THE INTERVIEW:

TERMINATION DATE: ________ TERMINATION WEEK: SUNDAY, ____ TO SATURDAY, ____

13TH WEEK: SUNDAY, ___ TO SATURDAY, ____

EMPLOYED AT TERMINATION? NO YES IF YES: JOB TITLE: ______________

I.D. 1-9

Our records indicate that you officially left the program on _______. I first want to ask you about the 13th week after you left the program; that is, the 7-day period starting on Sunday ____ and ending on Saturday ____, which was (last week/2 weeks ago/3 weeks ago). CARD "2" 10

1. Did you do any work for pay during that week?

NO . . 1

YES . . 2 P1 11

2. How much did you get paid for the work you did that week, before taxes? Include wages on all jobs, tips, overtime, and any work you may have done on the side.

___ ___ ___ ___ .00 DOLLARS ALL WEEK DK 9999 P2 12-15

RECORD NOTES FOR CALCULATIONS HERE:
HOURLY WAGE: ________
HOURS WORKED: ________
MONTHLY PAY: ________

3. Now I want to ask you about the entire 13 weeks starting with the week after you left the program. That is, from Sunday, ____ to Saturday, ___. Including the week we just talked about, how many weeks did you work at all for pay during the 13-week period?

00 01 02 03 04 05 06 07 08 09 10 11 12 13 WEEKS DK 99 P3 16-17

IF NOT EMPLOYED DURING ANY WEEK FOLLOWING TERMINATION, SKIP TO Q. 27, PAGE 5.

IF SOME EMPLOYMENT, BUT NOT DURING THE 13TH WEEK, SKIP TO TOP OF PAGE 4.

IF EMPLOYED DURING THE 13TH WEEK, CONTINUE, TOP OF NEXT PAGE.

Now I have a few questions about your job of (last week/2 weeks ago/3 weeks ago).

SKIP TO QUESTION 5 IF NOT EMPLOYED AT TERMINATION.

IF EMPLOYED AT TERMINATION, ASK:

4. Is this the same job you had when your program ended, back in (DATE)?

NO . 1 P4 18
YES . 2 -- **SKIP TO QUESTION 9**

5. How many weeks past (TERMINATION DATE) did you work on that job?

00 01 02 03 04 05 06 07 08 09 10 11 12 13 WEEKS DK 99 P5 19-20

6. What is the title of your job? (RECORD, CODE LATER)

CODE: ___ ___ ___ P6 21-23
TRAINING RELATED:
NO . 1 YES . 2 P7 24
DK . 9

IF TYPE OF WORK IS NOT CLEAR, PROBE FOR DUTIES
IF NO LONGER HAS THIS JOB, USE PAST TENSE FOR Q. 7-17

NOTES:

7. What kind of business or industry are you working for? SIC: ___ ___ P8 25-26
(e.g., bank, grocery, auto repair...)
RECORD AND CODE LATER:

8. How long ago did this job start? (OK TO RECORD DATE AND CODE LATER)

00 01 02 03 04 05 06 07 08 09 10 11 12 13 WEEKS DK 99 P9 27-28

9. How many hours per week do (did) you usually work on this job?

HOURS: ___ ___ PER WEEK. (98=98+) DK 99 P10 29-30

10. Is (was) that number of hours....

About what you want 1
More than you would like . 2 P11 31
or less than you would like . 3 DK 9

11. How much does (did) this job pay before deductions?

AMOUNT $ ___ ___ ___ ___ ___ . ___ ___ DK/NA 9999999 P12 32-38
PER: HOUR . 1 WEEK . 2 MONTH . 3 YEAR . . 4
OTHER . 5 SPECIFY:______________ DK 9 P13 39

DATA ENTRY

12. Does (did) this job offer......	NO	YES	DK		
a. Health insurance?	1	2	3	P14	40
b. A retirement program?	1	2	3	P15	41
c. Paid vacations?	1	2	3	P16	42
d. Higher pay for overtime	1	2	3	P17	43
e. A schedule of raises	1	2	3	P18	44

13. If you (had) stayed on this job two years, how likely would you be (have been) to get a promotion? Would you say . . .

Very . 1 somewhat . 2 a little bit . 3 or not at all likely . 4 P19 45

14. Assuming you (had) wanted to, how likely is it that you could keep (have kept) this job for the next year? Would you say it is . . .

Very . 1 somewhat . 2 a little bit . 3 or not at all likely . 4 P20 46

15. How well would you say your JTPA program prepared you for the work you are (were) doing? Would you say . . .

Extremely . 1 Very . 2 Moderately . 3 P21 47
Somewhat . 4 or Not at all well . 5 NA, DK 9

16. How much of the time on this job do (did) you use skills you learned during your JTPA enrollment? Would you say . . .

All of the time . . 1 most of the time . 2 P22 48
about half the time . 3 some of the time . . 4
seldom . . . . . . 5 or never . . . . . . 6 NA, DK 9

17. Overall, how satisfied are (were) you with your job? Would you say . . .

Extremely . 1 Very . 2 Moderately . 3 P23 49
Somewhat . 4 or Not at all satisfied . 5 DK 9

18. Since (TERMINATION DATE), have you had any jobs other than the one we have just been talking about?

NO . . . . 1 **SKIP TO Q 32.**
YES, THE TERMINATION JOB . . . . 2 **SKIP TO Q 32.** P24 50
YES, OTHER THAN TERMINATION JOB . 3 **ASK Q 20, NEXT PAGE**

DATA ENTRY

PAGE 4. START HERE IF EMPLOYED SOME, BUT NOT AT 13 WEEKS

****

Now we want to get a fuller record of the jobs you have had since (TERMINATION DATE). Lets begin with that date and work forward.

IF EMPLOYED AT TERMINATION, SAY:

My records show you were employed as a (TITLE) ON (TERMINATION DATE). Lets start there.

THEN ASK QUESTION 21

IF NOT EMPLOYED AT TERMINATION, ASK:

19. How many weeks after (TERMINATION DATE) was it until you found work?

00 01 02 03 04 05 06 07 08 09 10 11 12 13 DK 99 — P25 51-52

20. What was the title of that job?
(PROBE TITLE AND DUTIES UNTIL CLEAR)

____

CODE: _ _ _ — P26 53-55
TRAINING RELATED:
NO . 1 YES . 2 — P27 56
DK . 9

21. How many weeks did you work on that job?

00 01 02 03 04 05 06 07 08 09 10 11 12 13 DK 99 — P28 57-58

22. How many hours per week did you usually work on that job?

HOURS: ___ ___ PER WEEK. (98=98+) DK 99 — P29 59-60

23. How much did the job pay, before taxes?

PAY RATE: $ ___ ___ ___ ___ ___ . ___ ___ DK 9999999 — P30 61-17
PER: HOUR . 1 WEEK . 2 MONTH . 3 YEAR . 4 9
OTHER . 5 (SPECIFY:______________________________ DK 9 — P31 68

24. What is the main reason that job ended? (CODE INTO CLOSEST CATEGORY) — P32 69

FIRED 1
LAID OFF, CUTBACKS . . 2
END OF TEMPORARY JOB . 3
QUIT, HAD PROBLEMS . . 4
QUIT, DID NOT LIKE WORK . . 5
QUIT FOR BETTER JOB 6
OTHER 9:
____

DATA ENTRY

25. Have you had any other jobs since then?

NO 1

YES, THE 13 WEEK JOB . 2 P33 70

YES, OTHER JOB 3

26. How much would you say you earned from all jobs we have not already talked about?

___ ___ ___ ___ DOLLARS. 9999 DK P34 71-74

IF EMPLOYED CURRENTLY, SKIP TO QUESTION 32, NEXT PAGE.

BEGIN NEW CARD
I.D. 1-9
CARD "3" 10

IF NOT WORKING CURRENTLY, ASK QUESTION 27.

27. How serious a problem is it for you to be unemployed right now? Is it...

extremely serious . 1 very serious . . 2 P35 11

somwhat serious . . 3 or not too serious . 4 DK 9

28. Are you actively looking for work at this time?

NO . 1

YES . 2 DK 9 P36 12

29. OK. What are the main reasons you are not looking for work at this time? (RECORD. DO NOT CODE)

SKIP TO Q. 32, NEXT PAGE

30. Are you looking primarily for work in the area of your JTPA program, or for all types of work?

PROGRAM AREA ONLY . 1 PRIMARILY PROGRAM AREA . 2

ALL AREAS . 3 P37 13

31. How likely do you think it is that you will be able to find work within the next 2 months? Would you say . . .

Very . 1 somewhat . 2 a little bit . 3 P38 14

or not at all likely . 4 DK 9

DATA ENTRY

ASK ALL:

32. Are you, or any of your dependents, currently receiving public assistance payments of any type? (IF YES, PROBE FOR WELFARE VERSUS OTHER)

NO 1
YES, OTHER THAN WELFARE . . 2 — P39 15
YES, AFDC OR OTHER WELFARE . 3
BOTH 2 AND 3 . . 4

33. How much are you receiving this month?

___ ___ ___ ___ .00 DOLLARS — P40 16-19

34. How satisfied are you with the information and guidance you received from the staff at the JTPA office? Would you say . . .

Extremely . 1 Very . 2 Moderately . 3 — P41 20
Somewhat . 4 or Not at all satisfied . 5 DK 9

35. How satisfied are you with your JTPA program? (TRAINING OR PLACEMENT)

Extremely . 1 Very . 2 Moderately . 3 — P42 21
Somewhat . 4 or Not at all satisfied . 5 DK 9

36. How important is it to you that you have work which uses the skills you learned during your JTPA program? Would you say . . .

Extremely . 1 Very . 2 Moderately . 3 — P43 22
Somewhat . 4 or Not at all important . 5 DK 9

37. Overall, would you say that because of JTPA, your life has ...

improved a great deal . 1 improved a little . 2 not changed . 3 — P44 23
gotten a little worse . . 4 or gotten a lot worse . 5 DK 9

That completes our survey. Thanks very much for your cooperation. If there are any comments you'd like to add, please feel free to do so..... Thank-you.

EMPLOYER SURVEY, PART 1.

OUTCOME QUESTIONS FOR ALL EMPLOYERS

DATA ENTRY

PARTICIPANT ID	1-9
CARD 4	10
EMPLOYER ID	11-14

1. Is (PARTICIPANT) still working for you now? — E1 — 15

NO . 1 -- **SKIP TO TOP OF NEXT PAGE, QUESTION 6.**
YES . 2

2. What is the title of (PARTICIPANT'S) current job?

____

IF TYPE OF WORK IS NOT CLEAR FROM TITLE, ASK:
What are the main duties of this job? (RECORD:)

RECORD, CODE LATER
CODE : ___ ___ ___ — E2 — 16-18
TRAINING RELATED:
NO . 1
YES . 2 — E3 — 19
DK . 9

3. How many hours per week does (PARTICIPANT) usually work?

HOURS: ___ ___ PER WEEK. DK 99 — E4 — 20-21

4. How much does (PARTICIPANT) earn, before deductions?

AMOUNT: $ ___ ___ ___ ___ ___ . ___ ___ — E5 — 22-28
PER: HOUR . 1 WEEK . 2 MONTH . 3 YEAR . . 4
OTHER . 5 ______________________________ — E6 — 29

5. During the next year, would you say it is very likely, somewhat, a little bit, or not at all likely that (PARTICIPANT) will . . .

	VERY	SOME	LITTLE	NOT	DK		
a Be promoted	1	2	3	4	9	E7	30
b. Receive a raise	1	2	3	4	9	E8	31
c. Be fired	1	2	3	4	9	E9	32
d. Be laid off due to cutbacks	1	2	3	4	9	E10	33
e. Decide to quit	1	2	3	4	9	E11	34
f. Receive further training as part of the job	1	2	3	4	9	E12	35

SKIP ONE PAGE, TO QUESTION 10, QUESTIONS FOR ALL

DATA ENTRY

QUESTION 6. How long did (PARTICIPANT) work for you?

___ ___ WEEKS DK 99 E13 36-37

7. Did (PARTICIPANT) quit or did you let him(her) go?

QUIT 1 E14 38
FIRED 2
COMBINATION . 3

8. Do you know why?

PERSONAL PROBLEMS . 1
TROUBLE DOING WORK . 2
DISLIKED JOB . . . 3 E15 39
OTHER (EG MOVED) . . 4
NO, DON'T KNOW . . . 9

9. Could you tell me the main reason you let him(her) go?

BUSINESS CUTBACKS, PLAN TO REHIRE 1
BUSINESS CUTBACKS, NO PLAN TO REHIRE . . 2
END OF TEMPORARY OR SEASONAL JOB 3
FIRED: UNABLE TO DO THE WORK 4
FIRED: UNRELIABLE (MISSED WORK, LATE) . . 5
FIRED: ATTITUDE OR PERSONALITY PROBLEM . 6 E16 40
FIRED: NEGLIGENT OR DISHONEST 7
OTHER:________________________________ . 8

DATA ENTRY

QUESTIONS FOR ALL:

10. Would you please rate (PARTICIPANT) by comparing him(her) with the average workers you have hired or could hire for this same job. For each quality I read, please indicate whether (PARTICIPANT) is(was) much better ... a little better ... about the same ... a little worse ... or much worse than the average worker. First . . .

	MCH BET	LIT BET	SAME	LIT WRSE	MCH WRSE		
a. Learning new job skills	1	2	3	4	5	E17	41
b. Getting the work done quickly	1	2	3	4	5	E18	42
c. Following directions well	1	2	3	4	5	E19	43
d. Being willing to do extra work	1	2	3	4	5	E20	44
e. Working well independently	1	2	3	4	5	E21	45
f. Being enthusiastic on the job	1	2	3	4	5	E22	46
g. Getting along well with others	1	2	3	4	5	E23	47
h. Being honest	1	2	3	4	5	E24	48
i. Having the skills required for the job	1	2	3	4	5	E25	49
j. Being able to handle job stresses . .	1	2	3	4	5	E26	50
k. Keeping personal problems from interrupting work	1	2	3	4	5	E27	51

11. Using the same answer scale, how would you say (PARTICIPANT) compares to the average worker for . . .

	MCH BET	LIT BET	SAME	LIT LES	MCH LES		
a. Overall productivity as a worker	1	2	3	4	5	E28	52
b. Overall ease of supervision . . .	1	2	3	4	5	E29	53

12. Again, compared to the average worker, how fully trained was (PARTICIPANT) when you first hired him(her)?

MCH BET	LIT BET	SAME	LIT LES	MCH LES		
1	2	3	4	5	E30	54

IF THIS CONCLUDES THE INTERVIEW, THANK THEM AND ASK IF THEY HAVE ANY FINAL COMMENTS TO ADD

EMPLOYER SURVEY, PART 2.

QUESTIONS DESCRIBING THE FIRM

(These can be asked of all employers. Participating employers will normally have been asked this type of question during initial interviews)

DATA ENTRY

PARTICIPANT ID	1-9
CARD 5	10
EMPLOYER ID	11-14

Now I'd like to ask you just a few questions about your business and (PARTICIPANT'S) job.

13. First, are you a for-profit, private non-profit, or government employer?

PROFIT . 1 PRIVATE NON-PROFIT . 2 GOVERNMENT . 3 E31 15

14. What are the main products or services you provide?

RECORD:

SIC CODE: ___ ___ E32 16-17

15. How many employees do you have? (AT THIS BRANCH OR PLANT)

___ ___ ___ DK 999 E33 18-20

16. Over the past two years, has the volume of your business . . .

Grown a great deal . 1 Grown a little . 2 Remained stable . 3 E34 21

Declined a little . 4 or Declined a great deal . 5 DK 9

17. Does the volume of your business fluctuate widely on a seasonal basis, or because you do large contract work?

NO, IS STABLE 1 YES, CONTRACTS 3 E35 22

YES, SEASONAL 2 YES, BOTH . . 4 DK 9

18. When you hire a new (PARTICIPANT'S JOB TITLE), how long is it on the average until that job opens up again?

___ ___ MONTHS DK 99 E36 23-24

19. When you hire a (TITLE), do you normally hire someone who is already fully trained, partially trained, or not trained for this particular type of work?

FULLY . 1 PARTLY . 2 NOT TRAINED . 3 DK 8 NA 9 E37 25

DATA ENTRY

20. Either now or after a probationary period, does (PARTICIPANT'S) job include...

	NO	YES		
a. Health insurance?	1	2	E38	26
b. A retirement program?	1	2	E39	27
c. Paid vacations?	1	2	E40	28
d. Higher pay for overtime	1	2	E41	29
e. A schedule of raises	1	2	E42	30
f. Potential for promotion	1	2	E43	31

21. Are each of the following traits essential ... very important ... somewhat important ... or not important qualifications for (PARTICIPANT'S JOB)?

	ES'L	VERY	SOME	NOT	DK		
a. The ability to greet the public	1	2	3	4	5	E44	32
b. The ability to make complex decisions	1	2	3	4	5	E45	33
c. The ability to deal with the unexpected	1	2	3	4	5	E46	34
d. Formal training with this type of work	1	2	3	4	5	E47	35
e. Past experience with this type of work	1	2	3	4	5	E48	36
f. The ability to work without supervision	1	2	3	4	5	E49	37
g. The ability to do repetitive work well	1	2	3	4	5	E50	38
h. A strong educational background	1	2	3	4	5	E51	39

IF THIS ENDS THE SURVEY, THANK THE EMPLOYER AND ASK IF THEY HAVE ANY COMMENTS TO ADD.

ALTERNATIVELY, ADD QUESTIONS FOR PARTICIPATING EMPLOYERS ONLY

AND/OR ADD MARKETING OR OTHER QUESTIONS

EMPLOYER SURVEY, PART 3.

QUESTIONS FOR PARTICIPATING EMPLOYERS

DATA ENTRY

Since you have participated (are participating) in JTPA services, I'd like to ask you just a few more questions.

PARTICIPANT ID 1-9
CARD "6" 10
EMPLOYER ID 11-14

22. To date, how many JTPA participants has your organization hired?
____ ____ (RECORD NUMBER) DK 99 E52 15-16

23. On the whole, would you say the qualifications of JTPA applicants are. . . much better, a little better, about the same, a little worse, or much worse than the qualifications of non-JTPA applicants you get for the same jobs?
MUCH BETTER . 1 A LITTLE BETTER . 2 THE SAME . 3 E53 17
A LITTLE WORSE . 3 MUCH WORSE . 4 DK 9

24. How would you evaluate JTPA's administrative efficiency and responsiveness? Would you say they are . . .
Excellent . 1 Good . 2 Fair . 3 or Poor . 4 DK/NA 9 E54 18

25. On the same scale, what is your evaluation of the JTPA representative you have worked with most recently?
Excellent . 1 Good . 2 Fair . 3 or Poor . 4 DK/NA 9 E55 19

26. How would you evaluate how well JTPA screens applicants they send you?
Excellent . 1 Good . 2 Fair . 3 or Poor . 4 DK/NA 9 E56 20

27. On the same scale, what is your evaluation of the overall JTPA program, as you have experienced it?
Excellent . 1 Good . 2 Fair . 3 or Poor . 4 DK 9 E57 21

28. Do you plan to participate in the JTPA program again in the future?
NO . 1 UNDECIDED . 2 YES . 3 E58 22
WANT TO BUT PROBABLY CANNOT . 4

29. When you hired (PARTICIPANT) did you consider applications from JTPA referrals only or from other sources as well?

JTPA ONLY . . . . 1 SKIP TO QUESTION 31. DK 9 E59 23

OTHER (OR ONLY NON-JTPA . 2

DATA ENTRY

30. Did JTPA refer (PARTICIPANT) to you, or did you send him(her) to JTPA to check his(her) eligibility?

JTPA REFERRED . 1 — E60 24

EMPLOYER SENT . 2 DK 9

31. Did you learn that (PARTICIPANT was eligible before or after you had decided to hire him(her)?

BEFORE . 1 TENTATIVE BEFORE . 2 AFTER . 3 DK 9 — E61 25

32. How long was it until (PARTICIPANT) was fully trained for his/her job?

___ ___ WEEKS NOT YET OR NEVER . 99 DK . 98 — E62 26-27

33. During that training period, how many hours per week were devoted to training, on the average?

___ ___ HOURS PER WEEK DK 99 — E63 28-29

34. Approximately what percent of (PARTICIPANT'S) training involved written materials, courses, or other specialized training materials? (IF HESITANT, SUGGEST ROUNDING TO NEAREST 10%)

___ ___ ___ PERCENT DK 999 — E64 30-32

35. Approximately what percent of (PARTICIPANT'S) training would you say applied only to your firm, and would not carry over to some other employer? (IF HESITANT, PROMPT TO NEAREST 10%)

___ ___ ___ PERCENT DK 999 — E65 33-35

36. On the average, if you hired someone with no experience or training as a (PARTICIPANT'S JOB TITLE) how long would it be until they were fully trained?

NUMBER: ___ ___ — E66 36-37

OF: DAYS . 1 WEEKS . 2 MONTHS . 3 — E67 38

DATA ENTRY

37. Did you find that (PARTICIANT) required more, the same, or less training than others hired for the same job?

MORE . 1
SAME . 2
LESS . 3

E68 39

38. Could you estimate how much (more)(less) training, in percentage terms, such as 20% more or 20% less?

___ ___ ___ PERCENT E69 40-42

MORE . 1 LESS . 2 E70 43

39. Are your training methods for JTPA employees different in any way than training for others you hire for the same job?

NO . 1 YES . 2 DK 9 E71 44

40. There are various aspects to different JTPA programs which may or may not apply to you. For each I will read, would you please indicate whether it applies, and if so, whether you found it to be ... a major benefit ... a minor benefit ... neutral ... a minor cost ... or a major cost, of your participation in the JTPA program.

	MAJ BEN	MIN BEN	NEUTRAL	MIN CST	MAJ CST	NA/DK		
a. The wage reimbursement . . .	1	2	3	4	5	9	E72	45
b. JTPA screening of applicants .	1	2	3	4	5	9	E73	46
c. Reimbursement for courses during training	1	2	3	4	5	9	E74	47
d. Requirements or limits imposed by JTPA	1	2	3	4	5	9	E75	48
e. Enlarging or stabilizing your work force	1	2	3	4	5	9	E76	49
f. Paperwork for JTPA	1	2	3	4	5	9	E77	50
g. Knowing you are helping people who need it	1	2	3	4	5	9	E78	51
h. The risk of hiring someone very difficult to deal with	1	2	3	4	5	9	E79	52
i. Participation in the JTPA program, overall	1	2	3	4	5	9	E80	53

THANKS. THIS CONCLUDES THE INTERVIEW. DO YOU HAVE ANY FINAL COMMENTS TO ADD?

SUGGESTED MEASURES OF INDIVIDUAL TREATMENT

AGENCY INDIVIDUAL TREATMENT RECORD

I. PRE-JTPA JOB HISTORY (IF NOT IN PRESENT MIS)

1. First, I need to get a quick job history for the past 12 months. For any job you held since (ONE YEAR AGO), I need to get the job title, how long you had the job, how many hours per week you usually worked on the job, and what your rate of pay was.

	TITLE	WEEKS WORKED W/IN LAST 52	HOURS/WEEK ON AVERAGE	HOURLY WAGE ON AVERAGE
JOB A.	______________________	__ __	__ __	__ __ . __ __
JOB B.	______________________	__ __	__ __	__ __ . __ __
JOB C.	______________________	__ __	__ __	__ __ . __ __
JOB D.	______________________	__ __	__ __	__ __ . __ __
E. ALL OTHERS COMBINED:		__ __	__ __	__ __ . __ __

2. In the last 5 years, what is the longest full time job you have held?

____ ____ MONTHS

3. Have you, or any of your dependents, received public assistance payments of any type during the last 12 months? (IF YES, PROBE FOR WELFARE VERSUS OTHER)

NO . . . . 1 -- **SKIP TO TOP OF NEXT PAGE**
YES, NOT WELFARE 2
YES, AFDC OR OTHER WELFARE . 3
BOTH 2 AND 3 = ... 4

4. How many of the last 12 months did you receive any payments?

____ ____ MONTHS

5. How much did you receive, on the average?

____ ____ ____ DOLLARS

6. Are you receiving any payments currently?

NO . 1   YES . 2

7. RECORD THE MAJOR REASONS THIS PERSON NEEDS EMPLOYMENT HELP, BASED ON INTAKE CONVERSATIONS:

NOT IN LABOR FORCE, LONG TERM (STUDENT, DISPLACED HOMEMAKER)	NO 1	YES 2
TOOK TIME OUT OF LABOR FORCE, BY CHOICE	NO 1	YES 2
HAD A SKILL AREA, BUT TRIED TO CHANGE AREA, AND CAN'T FIND WORK	NO 1	YES 2
HAD A SKILL AREA, BUT LOST JOB AND CANNOT FIND NEW WORK IN AREA	NO 1	YES 2
HAS NEVER MANAGED TO FIND LONG TERM, OR FULL TIME WORK	NO 1	YES 2
HAS FOUND JOBS, BUT HAS DIFFICULTY KEEPING THEM	NO 1	YES 2
OTHER________________________	NO 1	YES 2
OTHER________________________	NO 1	YES 2

8. RECORD INDIVIDUAL'S ESTIMATED LEVEL OF NEED, UPON ENTRY, IN EACH AREA, USING THIS SCALE:

1 = EXTREME NEED, SERVICES REQUIRED. 2 = SUBSTANTIAL NEED, SERVICES USEFUL.
3 = MARGINAL NEED. 4 = NO NEED IN THIS AREA.

NEED FOR INFORMATION IN THE FOLLOWING AREAS:	EXTRM NEED	SUBST NEED	MARGINL NEED	NO NEED
A. OWN EMPLOYMENT AND TRAINING NEEDS AND GOALS . .	1	2	3	4
B. TRAINING OPTIONS AVAILABLE LOCALLY	1	2	3	4
C. EMPLOYMENT OPTIONS AVAILABLE LOCALLY	1	2	3	4
D. PROPER JOB SEARCH AND JOB APPLICATION METHODS.	1	2	3	4
LEVEL OF NEED IN THE FOLLOWING AREAS:				
E. BASIC LITERACY (INCLUDING ESL)	1	2	3	4
F. JOB SKILLS AND/OR CREDENTIALS	1	2	3	4
G. CREDIBLE WORK EXPERIENCE	1	2	3	4
H. MOTIVATION OR DIRECTION IN WORK LIFE	1	2	3	4
I. WORK HABITS, ON-THE-JOB MATURITY	1	2	3	4
J. OTHER:__________________	1	2	3	4
K. OTHER:__________________	1	2	3	4
L. OTHER:__________________	1	2	3	4

9. RECORD WHETHER PARTICIPANT BELONGS TO ANY TARGET GROUP OR OTHER SPECIAL GROUP IDENTIFIED FOR THIS ANALYSIS, BUT NOT RECORDED IN AGENCY MIS.

A. GROUP 1: ______________________ NO . 1 YES . 2
B. GROUP 2: ______________________ NO . 1 YES . 2
C. GROUP 3: ______________________ NO . 1 YES . 2
D. GROUP 4: ______________________ NO . 1 YES . 2

10. RECORD ANY OTHER INDIVIDUAL CHARACTERISISTICS FOR THE ANALYSIS BUT NOT INCLUDED IN MIS:

A. ON WORK RELEASE . . . . NO . 1 YES . 2
B. SINGLE PARENT . . . . NO . 1 YES . 2
C. EVER ENROLLED IN CETA OR JTPA BEFORE? . . . . . NO . 1 YES . 2
D. ______________________ RECORD CODE HERE
E. ______________________ RECORD CODE HERE
F. ______________________ RECORD CODE HERE

11. DID THE INDIVIDUAL ASK EXPLICITLY FOR A PARTICULAR TYPE OF SERVICE: (RECORD YES FOR ANY REQUESTED, NO FOR ALL OTHERS. IF NONE, RECORD NO ON ALL)

CLASSROOM SKILL TRAINING . . . . NO . 1 YES . 2
: AT A PARTICULAR SCHOOL . . . . NO . 1 YES . 2
: IN A PARTICULAR FIELD/COURSE . . . . NO . 1 YES . 2
: ALREADY ENROLLED THERE? . . . . NO . 1 YES . 2
OJT OR WEX . . . . NO . 1 YES . 2
: SENT BY AN EMPLOYER . . . . NO . 1 YES . 2
: EMPLOYER SAID HIRED IF ELIGIBLE? . . NO . 2 YES . 2
JOB SEARCH ASSISTANCE . . . . NO . 1 YES . 2
BASIC EDUCATION . . . . NO . 1 YES . 2
ANY FORM OF IMMEDIATE INCOME . . . . NO . 1 YES . 2
OTHER ______________________ NO . 1 YES . 2

12. DEGREE OF FINANCIAL NEED AT APPLICATION:

CURRENT EMERGENCY . 1   SERIOUS, MUST HAVE NEW INCOME SOON . 2

MODERATELY SERIOUS, CANNOT TRAIN WITHOUT INCOME . 3   NOT TOO SERIOUS . 4

13. AGENCY INTAKE RECOMMENDATION FOR THIS PARTICIPANT:

MOST INTENSIVE INTAKE NEEDED . 1   PARTIAL INTAKE NEEDED . 2

"FAST TRACK" INTAKE OK IN THIS CASE 3

14. SUMMARY OF INTAKE SERVICES: (MAKE NOTES AT LEFT DURING INTAKE, AND ENTER FINAL FIGURES AFTER INTAKE IS COMPLETE)

A. HOURS IN GROUP INTAKE ORIENTATION . . . . ____ ____

B. HOURS IN INFORMATION/MOTIVATION WORKSHOPS . ____ ____

C. HOURS IN JOB SEARCH WORKSHOPS . . . . . . ____ ____

D. HOURS OF INDIVIDUAL INTAKE COUNSELING . . . ____ ____

E. GIVEN SPECIFIC TESTS OR OTHER DIAGNOSIS? NO 1 YES 2

F. OTHER:______________________________ NO 1 YES 2

15. LIST MAJOR PROGRAM ACTIVITIES IN SEQUENCE BELOW, USING CODE NUMBERS IN USE FOR LOCAL MIS OR DEVELOPED FOR THIS RESEARCH. CODE WEEKS USING NUMBERS 1-52, SEQUENCED BY CALANDAR YEAR.

	ACTIVITY CODE	TRAINER/ PROVIDER CODE	START WEEK	SCHEDULED END WEEK	ACTUAL END WEEK	COM- PLETED
A.	___ ___ ___	___ ___ ___	___ ___	___ ___	___ ___	N 1 Y 2
B.	___ ___ ___	___ ___ ___	___ ___	___ ___	___ ___	N 1 Y 2
C.	___ ___ ___	___ ___ ___	___ ___	___ ___	___ ___	N 1 Y 2
D.	___ ___ ___	___ ___ ___	___ ___	___ ___	___ ___	N 1 Y 2

16. IF MORE THAN ONE ACTIVITY ABOVE, WAS THIS A PRE-PLANNED SEQUENCE?

NO . 1   YES . 2

17. HOW WELL WAS THE AGENCY ABLE TO "FIT" THE TREATMENT TO THE PARTICIPANT'S NEEDS (IN PROGRAM OFFICER'S OPINION):

EXCELLENT FIT 1   GOOD FIT 2   ACCEPTABLE FIT 3   POOR FIT 4   DK 9

18. LIST PROGRAM COSTS FOR THIS INDIVIDUAL'S PROGRAM, AS INDICATED BELOW. LEAVE BLANK EACH TYPE OF COST NOT APPLICABLE TO THIS PARTICIPANT. CATEGORIZE COSTS AS ARRANGED FOR THIS RESEARCH, NOT AS PER LOCAL BOOKKEEPING.

NATURE OF COST:	PLANNED TOTAL COST	PLANNED TOTAL COST	
A. CLASS TR. TUITION/FEES	___ ___ ___ ___	___ ___ ___ ___	DOLLARS
B. CLASS TR. LIVING STIPEND	___ ___ ___ ___	___ ___ ___ ___	DOLLARS
C. EMPLOYER WAGE REIMBURSEMENT	___ ___ ___ ___	___ ___ ___ ___	DOLLARS
D. OTHER TRAINING WITH OJT/WEX	___ ___ ___ ___	___ ___ ___ ___	DOLLARS
E. OTHER WORKSHOPS OR CLASSES	___ ___ ___ ___	___ ___ ___ ___	DOLLARS
F. SUPPORT SERVICES		___ ___ ___ ___	DOLLARS
G. OTHER:_______________________		___ ___ ___ ___	DOLLARS

---

19. 17. IF ENROLLED IN CLASSROOM TRAINING, ESTIMATE FOR TRAINING PERIOD:

A. MONTHLY INCOME FROM ASSISTANCE OTHER THAN JTPA: ___ ___ ___ ___ DOLLARS

B. MONTHLY INCOME FROM SELF, FAMILY MEMBERS, ETC.: ___ ___ ___ ___ DOLLARS

20. IF ENROLLED IN AN EMPLOYER BASED TREATMENT, CODE FROM CONTRACT:

A. PLANNED STARTING WAGE RATE DURING CONTRACT $ ___ ___ . ___ ___

B. PERCENTAGE OF WAGE REIMBURSED BY JTPA . . . . . ___ ___ %

C. DOES CONTRACT INCLUDE A DELAYED INCENTIVE PAYMENT? NO . 1 YES . 2

NOTE: OTHER MEASURES SPECIFIC TO PROGRAM ACTIVITY OR SERVICE MAY BE ADDED, OR MAY BE APPENDED FROM OTHER DATA SOURCES.

21. RECORD ALL THAT APPLY TO THIS PARTICIPANT'S EXIT:

A. PLACED IN ADMINISTRATIVE HOLD/INFORMAL JOB SEARCH STATUS. NO 1 YES 2

IF YES: HOW MANY WEEKS? ____ ____ WEEKS

B. GIVEN JOB SEARCH ASSISTANCE:	WORKSHOP	NO 1	YES 2
	JOB CLUB PHONE BANK . . .	NO 1	YES 2
	INDIVIDUAL COUNSELING . .	NO 1	YES 2
	SPECIFIC JOB REFERRAL(S) .	NO 1	YES 2

C. IDENTIFIED AS OUT OF THE LABOR FORCE AT TERMINATION . . NO 1 YES 2

22. IF EMPLOYED AT TERMINATION:

A. RECORD TERMINATION EMPLOYER AND I.D. ON IDENTIFIER SHEET

B. HOW MANY WEEKS BEFORE THE OFFICIAL TERMINATION DATE DID THE PARTICIPANT BEGIN THE JOB?

___ ___ WEEKS

C. IS THE TERMINATION EMPLOYER ALSO THE PARTICIPATING EMPLOYER FOR THIS PARTICIPANT?

NO . 1 YES . 2

D. IS THIS POSITION TRAINING RELATED? NO . 1 YES . 2 DK . 9

E. IF WAGE AND HOURS PER WEEK ARE NOT INCLUDED IN MIS, THEY SHOULD BE ADDED TO THIS FORM.

DESCRIPTIONS, NAMES, AND MEASUREMENT TYPES FOR MEASURES INCLUDED IN APPENDIX E

INTRODUCTORY NOTE

The purpose of this exhibit is to help identify and organize measures in the Appendix E instruments for analysis. Each measure is described and then located via the variable names suggested in the right hand columns of Appendix E measurement instruments. For example, whether a participant is still with the termination employer at follow-up is measured by either P4 or E1. In addition, the measurement type for each variable is listed as a guide to the selection of appropriate statistics for analysis.

Variable names consisting of the letter P followed by numbers refer to the participant interview. Names beginning with E refer to the employer survey, and those beginning with AT refer to the suggested Agency Treatment form. Some names follow none of these conventions, but are full words or abreviated words. A few of these are named on the Master Identifier Sheet whih precedes the participant interview form. However most names of this type refer to variables which do not exist on one of the surveys, but instead must be constructed from other measures. In these cases, instructions for creating these variables are included in the exhibit.

VARIABLE DESCRIPTION	VARIABLE NAME IN SURVEYS	MEASUREMENT TYPE
Identifiers		
Participant I.D.	PARTIC	NOMINAL
SDA I.D.	SDA	NOMINAL
Subcontractor I.D	SUB	NOMINAL
Program Officer I.D.	OFFICER	NOMINAL
Classroom Trainer I.D.	CTID	NOMINAL
Participating Employer I.D.	PEMPID	NOMINAL
Termination Employer I.D.	TEMPID	NOMINAL
Training Field	FIELD	NOMINAL
Eligibility Date	YEAR1, WEEK1	ORDERED
CONSTRUCT ELIGDATE = ((YEAR1-6)*52)+WEEK1		
Enrollment Date	YEAR2, WEEK2	ORDERED
CONSTRUCT ENRLDATE = ((YEAR2-6)*52)+WEEK2		
Termination Date	YEAR3, WEEK3	ORDERED
CONSTRUCT TERMDATE = ((YEAR3-6)*52)+WEEK3		
Follow-up Interview Date	YEAR4, WEEK3	ORDERED
CONSTRUCT FOLWDATE = ((YEAR4-6)*52)+WEEK4		
Interviewer I.D.	INTID	NOMINAL

VARIABLE DESCRIPTION	VARIABLE NAME IN SURVEYS	MEASUREMENT TYPE
** level of welfare dependence, including:		
-- Whether receiving public assistance at follow-up.	P39	DICHOT.
-- Monthly rate of public assistance at follow-up. .	P40	ORDERED
-- Pre-to post-program change in public assistance status	P39,AT3,AT6	NOMINAL
-- Pre- to post-program level of public assistance CONSTRUCT PRE-PROGRAM LEVEL = AT5*AT4/12	P40,AT4,AT5	ORDERED
3. Measures of skill transfer and utilization.		
** Whether employment is in a training-related field . . CONSTRUCT = YES IF E3=2; IF P7=2; IF P4=2 AND AT22D=2; OR E1=2 AND AT22D=2	P7,P4,E1,E3 AT22D	DICHOT.
** Proportion of work using skills from training. . . .	P22	ORDERED
4. Measures of job quality, indicating "primary" versus "secondary" labor market jobs.		
** Benefits (medical, retirement, vacations, overtime)	P14 TO P18	DICHOT.
and schedule of raises)	E38 TO E42	DICHOT.
** Likelihood of layoffs	P20,E9,E10	ORDERED
** Likelihood of receiving a raise	E8	ORDERED
** Likelihood of being promoted	P19,E7,E43	DICH/ORD
** Amount of work adequate?	P11	ORDERED
** Likelihood of receiving further training from employer	E12	ORDERED
5. Measures characterizing those not employed or not retaining termination jobs.		
** Why termination job was lost or left	P32,E14-E16	NOMINAL
** Participant in the labor force at follow-up?	P36,P38	DICHOT.
6. Subjective orientations of participants.		
** Intention to make use of the JTPA intervention . . .	P43,P37	ORDERED
** Evaluations of JTPA program services	P41,P42,P44	ORDERED
** Evaluation of post-program job	P23	ORDED

VARIABLE DESCRIPTION	VARIABLE NAME IN SURVEYS	MEASUREMENT TYPE
Participant Outcome Measures		
1. Post-Program Performance Requirements.		
** Employed during 13th week?	P1	DICHOT
** Earnings during 13th week	P2	ORDERED
** Number of first 13 weeks had any employment	P3	ORDERED
2. Other Measures Core to the JTPA Mandate.		
**** Employment and employment intensity, including:**		
-- Still with termination employer?	P4, E1	DICHOT.
-- IF NOT: number of weeks retained	P5, E13	ORDERED
-- Still with participating employer?	P4,AT22C,E1	DICHOT.
CONSTRUCT YES IF P4 (OR E1) = 2 AND AT22C = 2 .		
-- Average hours/week employed on follow-up job.	P10,E4	ORDERED
-- Pre- to post-program change in hours per week	HOURCHNG	ORDERED
CONSTRUCT = P10 (OR E4) - PREHOURS (SEE BELOW)		
-- Pre- to post-program change in proportion of weeks employed	WEEKCHNG	ORDERED
CONSTRUCT = P3/13 - PREWEEKS (SEE BELOW)		
**** Income, including:**		
-- Hourly wage rate at follow-up.	HOURLY	ORDERED
CONSTRUCT: IF P13 = 1, HOURLY = P12 IF P13 = 2, HOURLY = P12/P10 IF P13 = 3, HOURLY = P12/(P10*4.3) IF P13 = 4, HOURLY = P12/(P10*52)		
OR: SAME AS ABOVE, USING E4, E5, AND E6.		
-- Monthly earnings at follow-up	MONTHLY	ORDERED
CONSTRUCT: MONTHLY = HOURLY * P10 (OR E4) * 4.3		
-- Total earnings from termination to follow-up	TOTEARN	ORDERED
CONSTRUCT: TOTEARN = HOURLY * P10 * P9 + PARALLEL MEASURES FOR JOB HISTORY, USING P28-P31) + P34		
-- Pre- to post-program change in hourly wage	HRLYCHNG	ORDERED
CONSTRUCT = HOURLY - PREHRLY (SEE BELOW)		

VARIABLE DESCRIPTION	VARIABLE NAME IN SURVEYS	MEASUREMENT TYPE
Outcomes for Termination Employers		
Performance on the job, compared to non-JTPA workers		
** Skill level.		
-- How fully trained when hired.	E30,E68	ORDERED
-- Skill level at follow-up	E25	ORDERED
** Job performance		
-- Overall productivity	E28	ORDERED
-- Getting work done quickly	E18	ORDERED
-- working well independently	E21	ORDERED
** Supervision ease and work habits		
-- Overall ease of supervision	E29	ORDERED
-- Following directions well	E19	ORDERED
-- Being willing to do extra work	E20	ORDERED
-- Being enthusiastic on the job	E22	ORDERED
-- Being honest and reliable	E24	ORDERED
** Personal adjustment on the job		
-- Getting along well with others	E23	ORDRED
-- Being able to handle job stresses	E26	ORDERED
-- Keeping personal life from hurting work	E27	ORDERED
Outcomes for Participating Employers, Measured Through Agency Records		
** Total wage reimbursement for this contract	AT18C	ORDERED
** Proportion of wages redimbursed	AT20B	ORDERED
** Time period of reimbursement	AT15	ORDERED
** Amount spent on other training with OJT or WEX	AT18D,E	ORDERED

VARIABLE DESCRIPTION	VARIABLE NAME IN SURVEYS	MEASUREMENT TYPE
Outcomes for Participating Employers, Measured Through Follow-up Surveys		
1. Measures of Training Provided		
** Number of weeks until full productivity	E62	ORDERED
** Intensity (number of hours/week) of training	E63	ORDERED
** total amount of training (E62*E63)	E62, E63	ORDERED
** Length of training compared to non-JTPA workers (TRANSFORM E66,E67 TO WEEKS AND SUBTRACT E62)	E62,E66,E67	ORDERED
** Amount of training compared to non-JTPA workers	E68-E70	ORDERED
** Use different methods to train JTPA workers?	E71	DICHOT.
** Use formal training methods?	E64	ORDERED
** Proportion of training non-transferable	E65	ORDERED
2. Measures of Employer Cost Via Risk Taking		
** How much hiring power assumed by employer	E59,E60,E61	DICHOT.
** Perceived risk of hiring a problem case	E79	ORDERED
3. Measures of Employer Cost Via Perceived Constraint		
** Perception of requirements or limits imposed by JTPA	E75	ORDERED
** Perception of JTPA paperwork requirements	E77	ORDERED
4. Employer Perceptions of JTPA and of Costs or Benefits of Participation		
** Evaluations of the JTPA program and service providing agencies.		
-- Administrative efficiency and responsiveness	E54	ORDERED
-- JTPA representative working with the employer	E55	ORDERED
-- JTPA screening of applicants sent to employers	E56	ORDERED
-- Qualifications of JTPA referrals, on the whole	E53	ORDERED
-- The overall JTPA program	E57	ORDERED
** Intention to participate in JTPA again	E58	NOMINAL
** Total frequency have hired JTPA participants	E52	ORDERED

VARIABLE DESCRIPTION	VARIABLE NAME IN SURVEYS	MEASUREMENT TYPE
** Perceptions of various aspects of participation as costly or beneficial		
-- The wage reimbursement	E72	ORDERED
-- JTPA screeming of applicants	E73	ORDERED
-- Reimbursement for courses during training	E74	ORDERED
-- Enlarging or stabilizing employer's work force	E76	ORDERED
-- The feeling of helping those with need	E78	ORDERED
-- Program participation, overall	E80	ORDERED

Pre-Program Measures Used to Construct Change Outcomes

NOTE: THESE VARIABLES ARE CONSTRUCTED FROM QUESTION 1 ON THE AGENCY INDIVIDUAL TREATMENT RECORD, WHICH USES A CONDENSED FORMAT. TO FACILITATE TRANSFORMATION INSTRUCTIONS, THE FORMAT IS SHOWN BELOW, ALONG WITH THE A NAMING CONVENTION USED BELOW. VARIABLE NAMES ARE SHOWN IN BOLD PRINT.

TITLE	WEEKS WORKED W/IN LAST 52	HOURS/WEEK ON AVERAGE	HOURLY WAGE ON AVERAGE
JOB A.______________________	**AT1A1**	**AT1A2**	**AT1A3**
JOB B.______________________	**AT1B1**	**AT1B2**	**AT1B3**

VARIABLE DESCRIPTION	VARIABLE NAME IN SURVEYS	MEASUREMENT TYPE
** Proportion of 52 pre-program weeks employed CONSTRUCT PREWEEK = AT1A1 + AT1B1 ... + AT1E1/52	PREWEEK	ORDERED
** Average hours per week while employed prior to program CONSTRUCT PREHOURS = AT1A1*AT1A2...+ AT1E1*AT1E2/PREWEEK*52	PREHOURS	ORDERED
** Average hourly wage during 52 pre-program weeks CONSTRUCT PREHRLY = AT1A1*AT1A2*AT1A3... + AT1E1*AT1E2*AT1E3/PREWEEK*52*PREHOURS	PREHRLY	ORDERED
** Average Monthly income during 52 pre-program weeks CONSTRUCT PREMONLY = PREHRLY * PREHOURS * PREWEEK * 4.3	PREMONLY	ORDERED

VARIABLE DESCRIPTION	VARIABLE NAME IN SURVEYS	MEASUREMENT TYPE

Agency Individual Treatment Program Variants

** Need for Screening, Intake Services, and Treatment		
-- Why participant applied to JTPA (series)	AT7,AT13	DICHOT.
-- Participant information needs (series)	AT8	ORDERED
-- Participant employment blocks (series)	AT8	ORDERED
-- Target or status information not in MIS (series)	AT9, AT10	DICHOT.
-- Requests for services by participant (series) .	AT11	DICHOT.
** Intake Services		
-- Intake intensity recommended	AT13	ORDERED
-- Intake route (selected items)	AT11	DICHOT.
-- Intake services provided (series)	AT14	ORDERED
** Basic Program Activities		
-- What service(s) enrolled in	AT15	NOMINAL
-- Planned and actual lengths of (each) activity .	AT15	ORDERED
-- Whether (each) activity was completed	AT15	DICHOT.
-- If more than one, are they in planned sequence?	AT16	DICHOT.
-- How well is the treatment judged to fit the need?	AT17	ORDERED
-- Planned and actual costs of each program activity	AT18	ORDERED
-- IF CT: sources of income	AT18B,AT19	ORDERED
-- IF EMPLOYER BASED: wage, reimubursement rate, and contract type	AT20	ORDERED
** Classroom Trainer Characteristics		
-- No ready made measures included in this guide		
** Participating Employer Characteristics		
-- The employer's size	E33	ORDERED
-- The employer's growth rate	E34	ORDERED
-- The typical turnover rate	E36	ORDERED
-- Industrial sector, and whether public or private.	E31, E32	NOMINAL
-- The quality and complexity of the job (series) .	E44 TO E51	ORDERED

VARIABLE DESCRIPTION	VARIABLE NAME IN SURVEYS	MEASUREMENT TYPE
** Ancillary Services		
-- Cost of classroom training added to OJT or WEX	AT18D	ORDERED
-- Cost of other workshops not coded as activities	AT18E	ORDERED
-- Cost of support services	AT18F	ORDERED
** Termination Data, all		
-- Placed in administrative hold, and how long . .	AT21A	ORDERED
-- Job search services provided	AT21B	DICHOT.
-- Out of labor force at termination?	AT21C	DICHOT.
** Termination Data if Employed		
-- Number of weeks before termination date got job	AT22B	ORDERED
-- Is termination employer also participating empl?	AT22C	DICHOT
-- Is the position training-related?	AT22D	
-- Employer ID, wages, hours per week, etc.	ASSUMED IN MIS	